

(12) UK Patent Application (19) GB (11) 2 390 779 (13) A

(43) Date of A Publication 14.01.2004

(21) Application No: 0216239.4

(22) Date of Filing: 12.07.2002

(71) Applicant(s):
Fujitsu Limited
 (Incorporated in Japan)
 1-1 Kamikodanaka 4-chome, Nakahara-ku,
 Kawasaki-shi, Kanagawa 211-8588, Japan

(72) Inventor(s):
Abedi Saied

(74) Agent and/or Address for Service:
Haseltine Lake & Co
 Imperial House, 15-19 Kingsway,
 LONDON, WC2B 6UD, United Kingdom

(51) INT CL⁷:
H04L 12/56, H04Q 11/04

(52) UK CL (Edition W):
H4L LDGP L201 L203 L205 L213
H4K KTKX

(56) Documents Cited:
GB 2361147 A **EP 1227626 A2**
WO 2001/063855 A1 **CN 001347253 A**

(58) Field of Search:
 UK CL (Edition V) **H4K, H4L**
 INT CL⁷ **H04L, H04Q**
 Other: **ONLINE: EPODOC, WPI, JAPIO**

(54) Abstract Title: **Packet scheduling**

(57) Packet scheduling apparatus 10 schedules packets of data for transmission from a transmitter via at least one channel to a plurality of receivers. Using conventional scheduling methods, the performance aspects of a scheduling solution, i.e. data throughput, delay, fairness of packet delivery and quality of service, are generally unbalanced, e.g. round-robin solutions are fair but with poor throughput while maximum carrier-to-interference (max C/I) solutions have good throughput but at the expense of fairness. The invention provides a packet scheduler which selects a scheduling solution whose performance aspects are best in relation to a balance of performance aspects set by an operator. A plurality of candidate scheduling solutions η_1 - η_n are generated by a candidate solution generating unit 12 (fig. 4). The solutions may include round-robin and max C/I solutions. At least one candidate solution is generated randomly and each candidate solution specifies at least the receiver(s) to which packets are to be transmitted in a scheduling instant under consideration. A best solution selecting unit 14 compares the generated candidate solutions, e.g. by comparing preselected fitness measures of each solution, and based on the comparison results, selects a best one of the candidate solutions η_{best} to use to transmit packets in the scheduling instant. Applications in WWW browsing, W-CDMA, IS95, TDMA, WDMA, FDMA, SDMA.

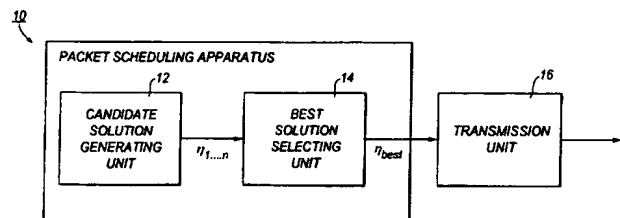
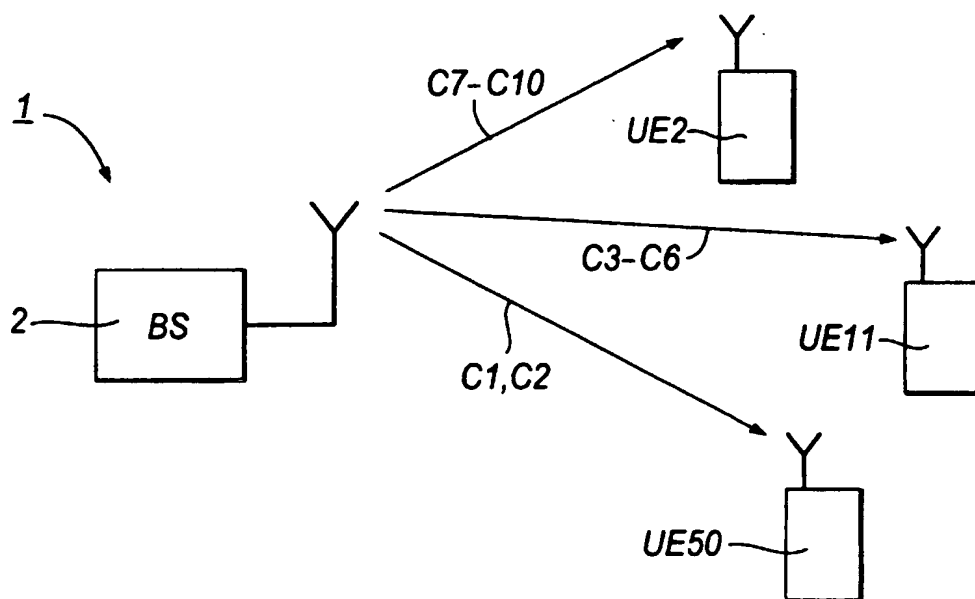


Fig.3

THIS PAGE BLANK (P. 00)

*Fig.1*

	TTI 1	TTI 2	TTI 3	TTI 4	TTI 5	TTI 6	TTI 7	TTI 8	TTI 9
Node B 1 Channel Code 1	UE50 Packet 1	UE1 Packet 1	UE3 Packet 1	UE23 Packet 1	UE4 Packet 2	UE6 Packet 1	UE1 Packet 1	UE23 Packet 1	UE7 Packet 1
Node B 1 Channel Code 2	UE50 Packet 2	UE50 Packet 3	UE4 Packet 1	UE1 Packet 2	UE7 Packet 1	UE50 Packet 2	UE4 Packet 1	UE1 Packet 2	UE17 Packet 1
Node B 1 Channel Code 3	UE11 Packet 1	UE50 Packet 4	UE1 Packet 1	UE4 Packet 2	UE23 Packet 2	UE2 Packet 8	UE11 Packet 4	UE15 Packet 1	UE23 Packet 2
Node B 1 Channel Code 4	UE11 Packet 2	UE50 Packet 5	UE50 Packet 10	UE11 Packet 8	UE16 Packet 2	UE9 Packet 2	UE40 Packet 2	UE34 Packet 1	UE9 Packet 3
Node B 1 Channel Code 5	UE11 Packet 3	UE2 Packet 5	UE16 Packet 1	UE9 Packet 1	UE11 Packet 3	UE24 Packet 2	UE43 Packet 2	UE16 Packet 1	UE11 Packet 10
Node B 1 Channel Code 6	UE11 Packet 4	UE11 Packet 5	UE2 Packet 6	UE1 Packet 3	UE4 Packet 1	UE11 Packet 9	UE40 Packet 3	UE11 Packet 5	UE4 Packet 1
Node B 1 Channel Code 7	UE2 Packet 1	UE50 Packet 6	UE50 Packet 11	UE3 Packet 2	UE24 Packet 1	UE2 Packet 8	UE38 Packet 1	UE38 Packet 5	UE50 Packet 14
Node B 1 Channel Code 8	UE2 Packet 2	UE50 Packet 7	UE50 Packet 12	UE11 Packet 9	UE43 Packet 1	UE4 Packet 3	UE38 Packet 2	UE3 Packet 2	UE14 Packet 1
Node B 1 Channel Code 9	UE2 Packet 3	UE50 Packet 8	UE11 Packet 6	UE40 Packet 1	UE50 Packet 13	UE3 Packet 1	UE38 Packet 3	UE40 Packet 1	UE43 Packet 2
Node B 1 Channel Code 10	UE2 Packet 4	UE50 Packet 9	UE11 Packet 7	UE50 Packet 13	UE2 Packet 7	UE43 Packet 2	UE38 Packet 4	UE50 Packet 14	UE2 Packet 9

Fig.2

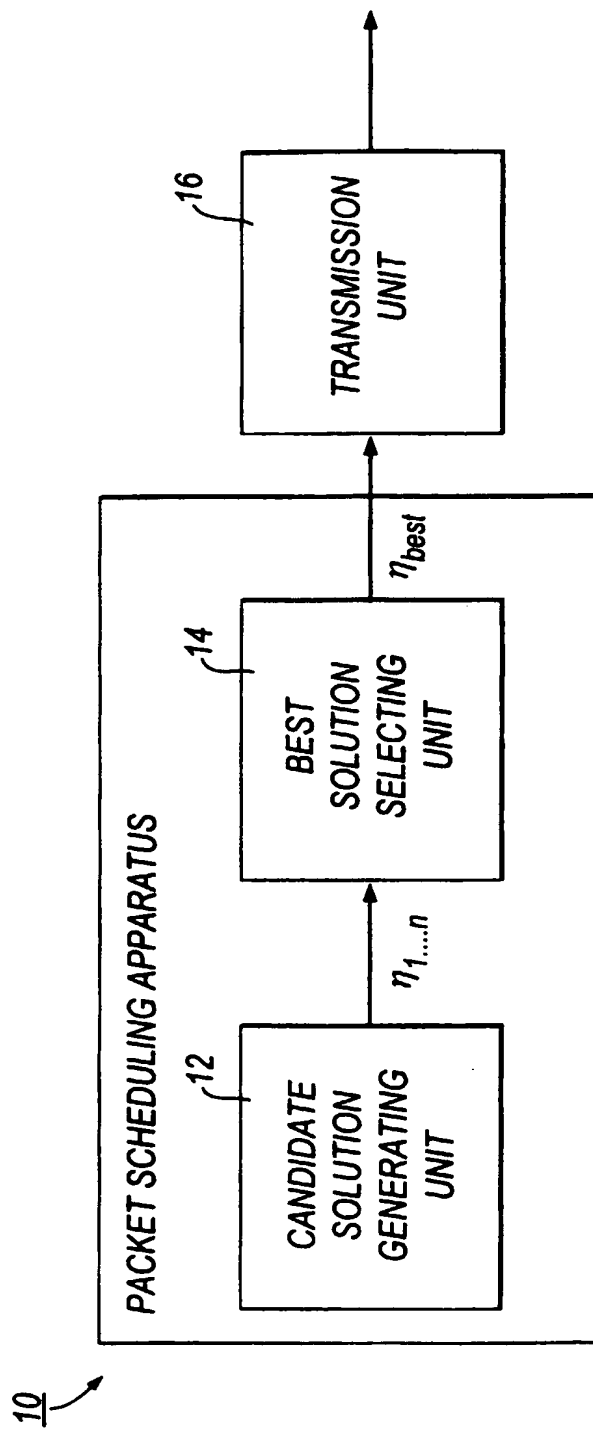


Fig.3

4/23

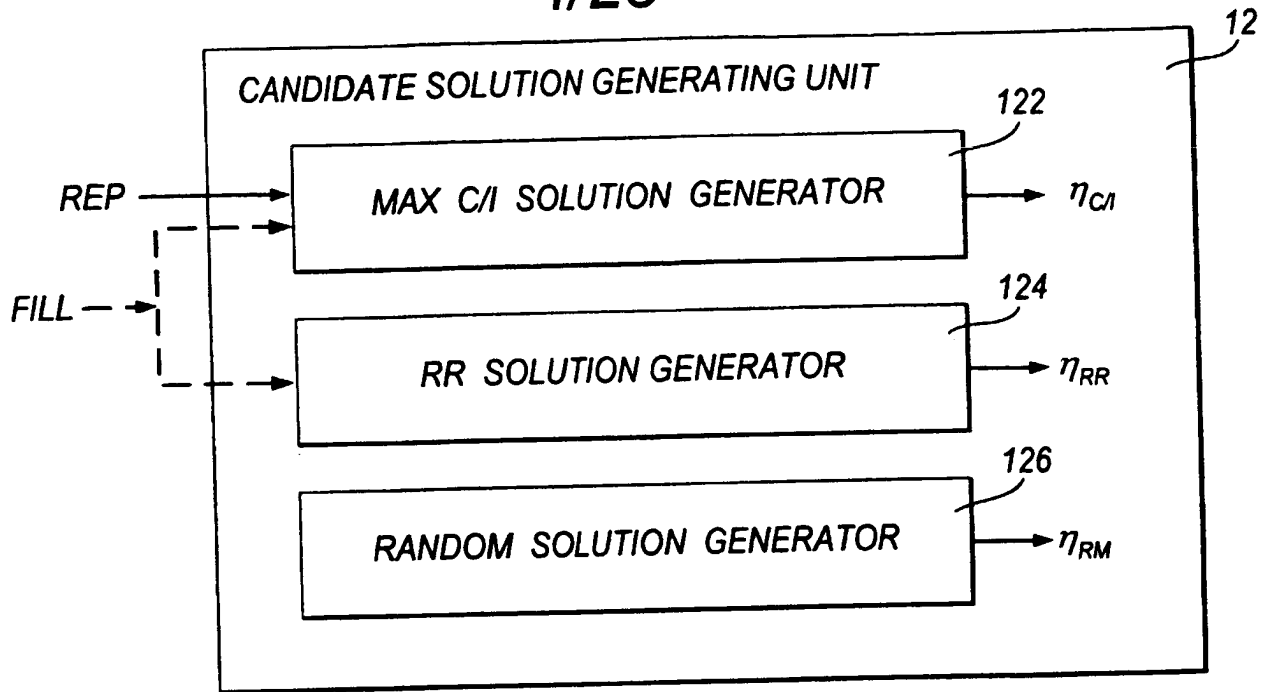


Fig.4

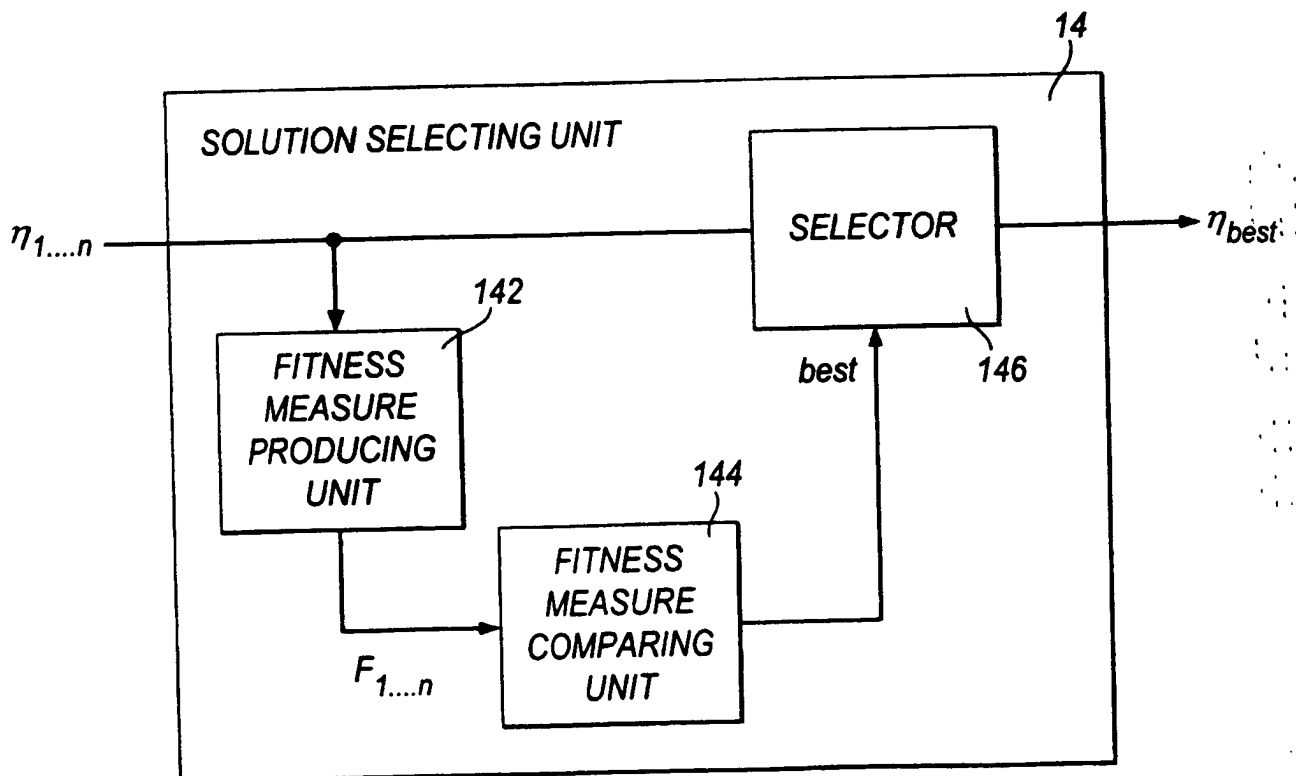


Fig.5

5/23

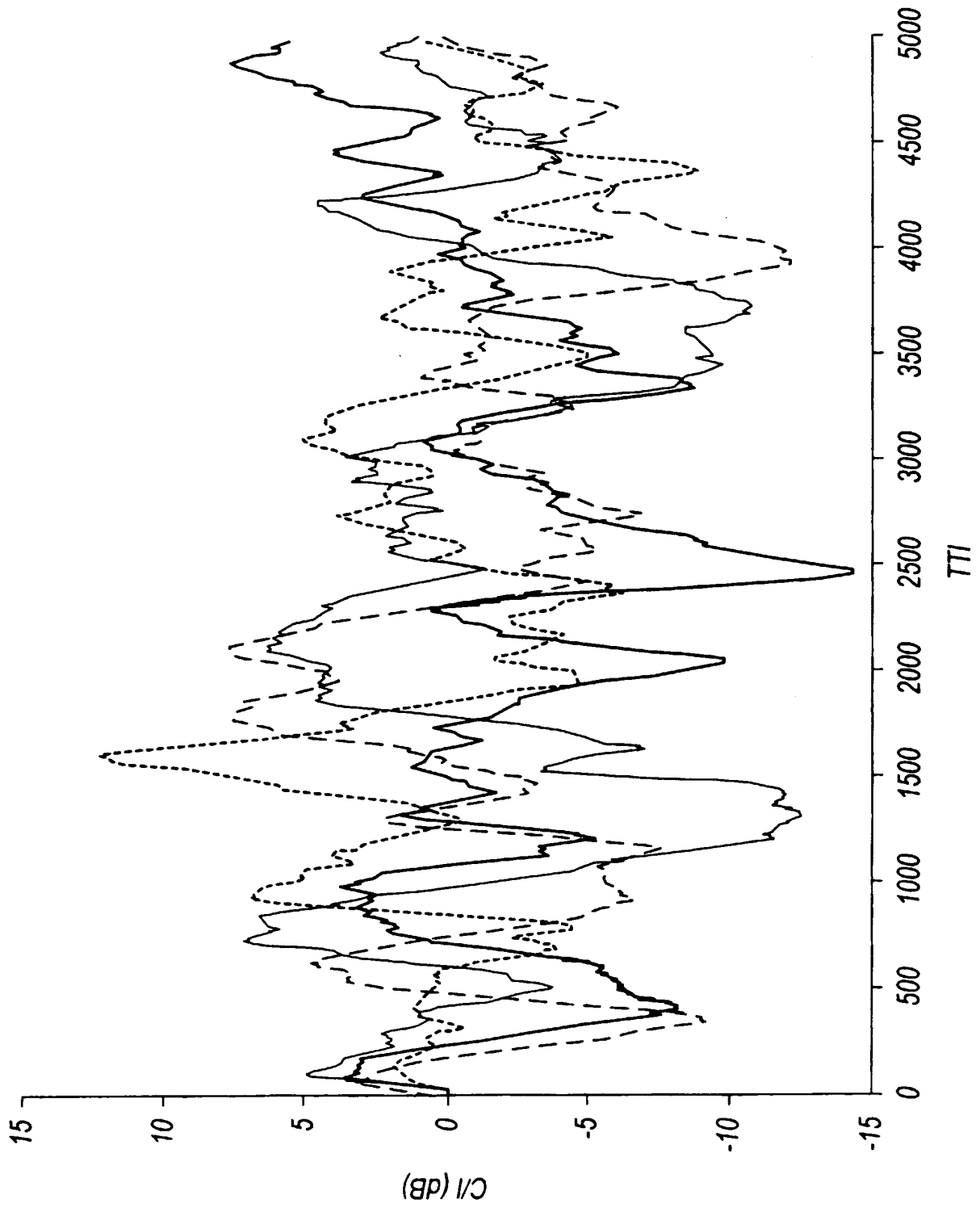


Fig.6

6/23

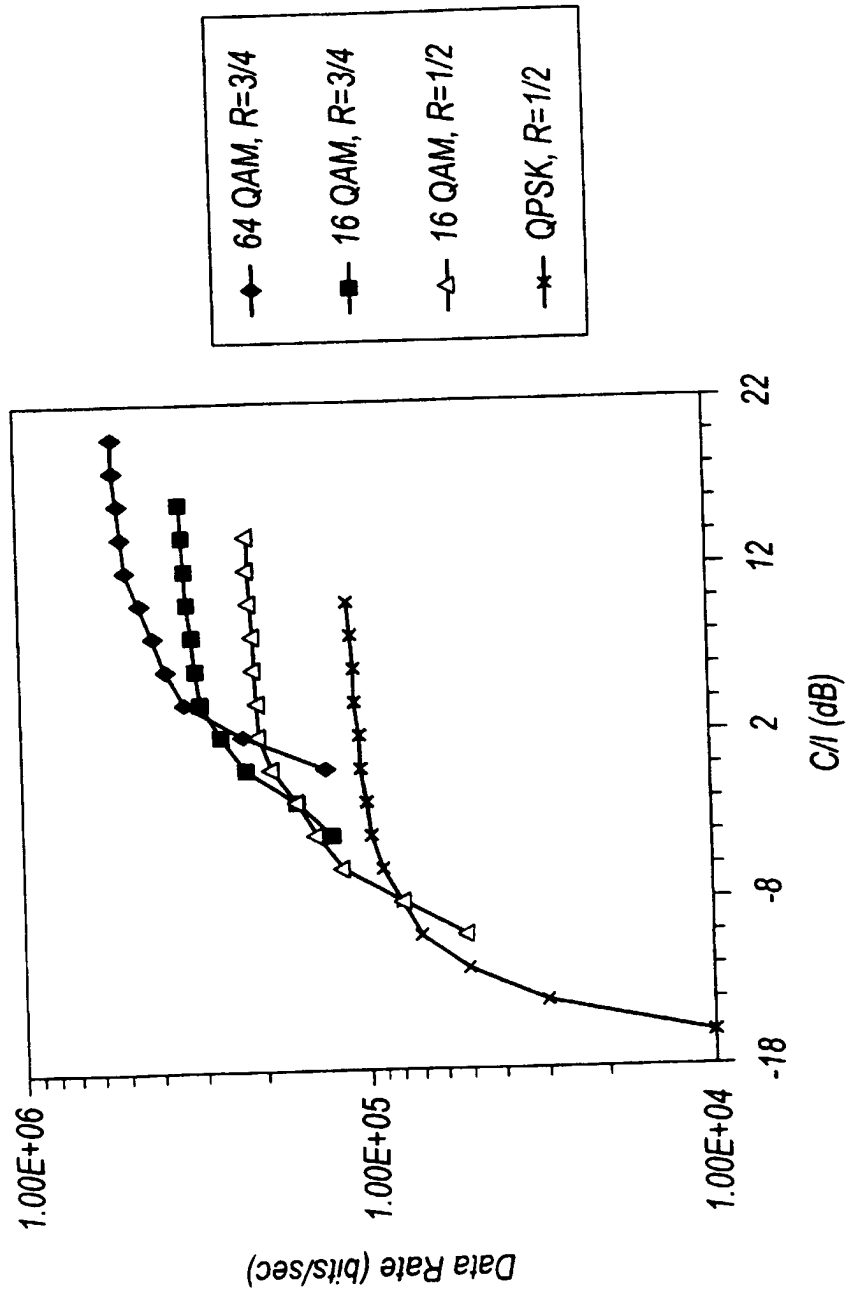


Fig. 7

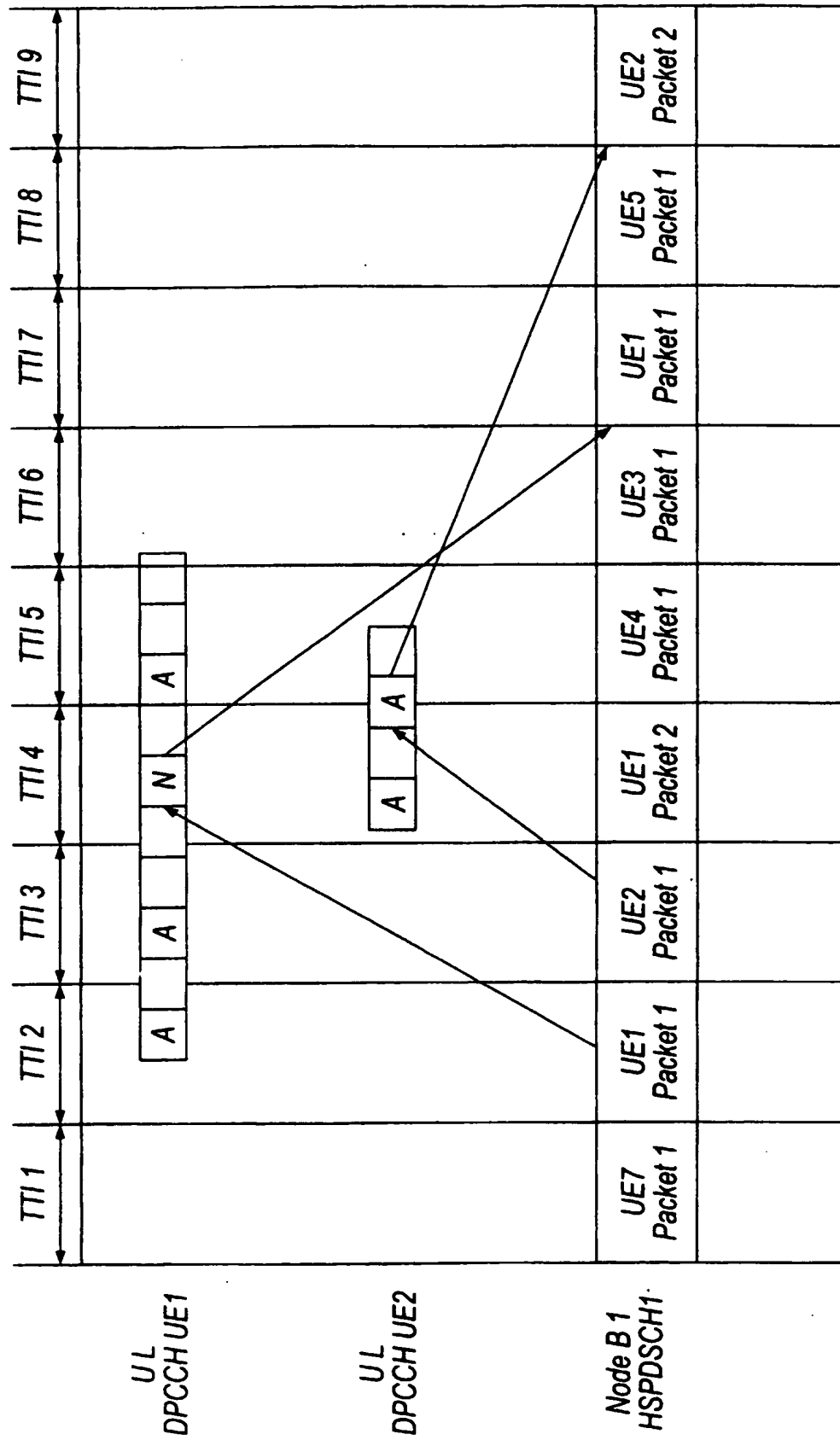


Fig.8

000000

8/23

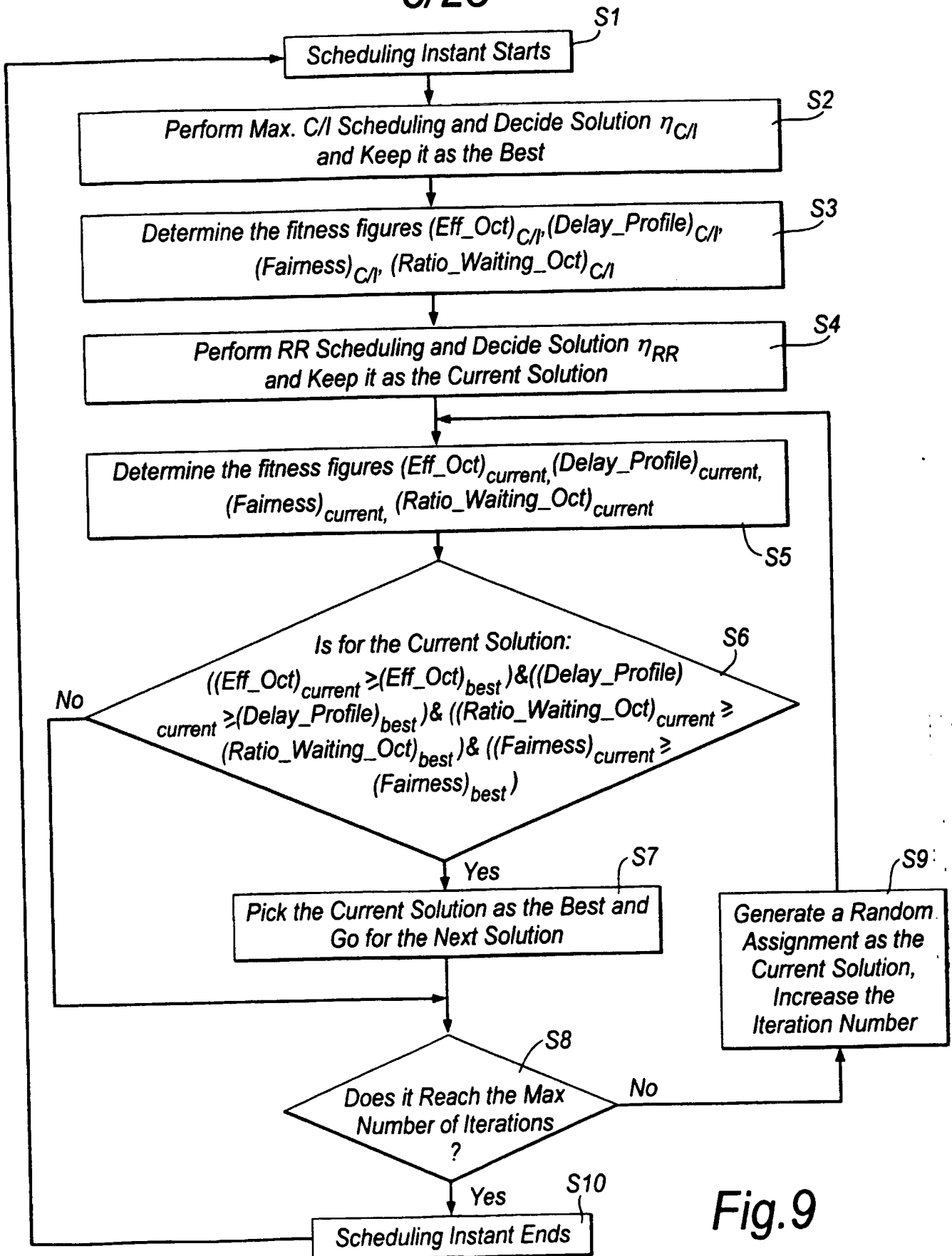
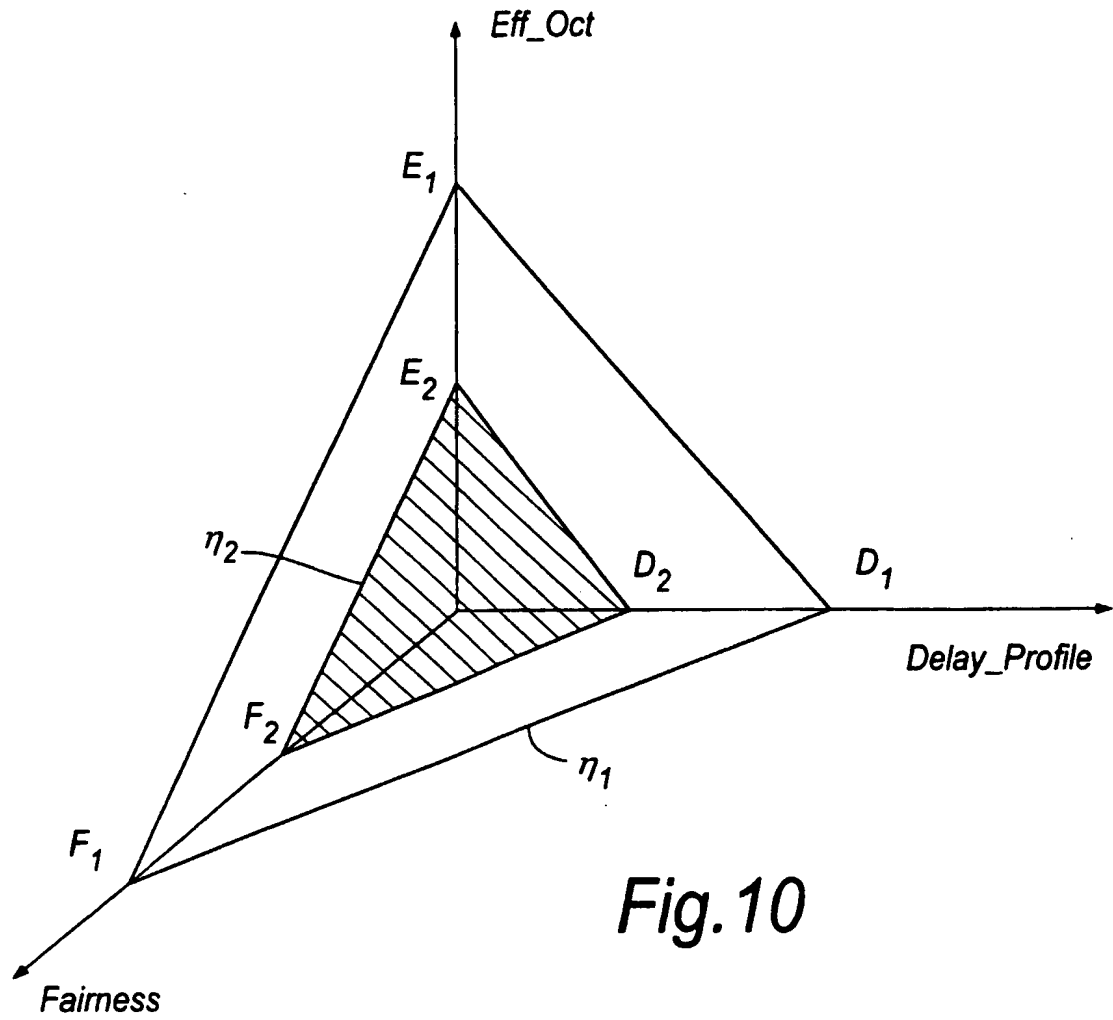
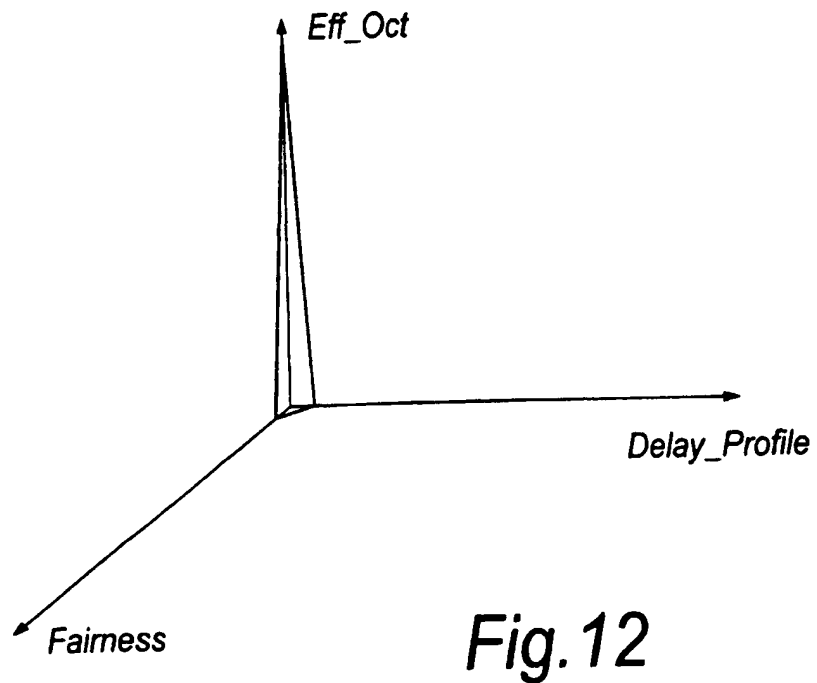
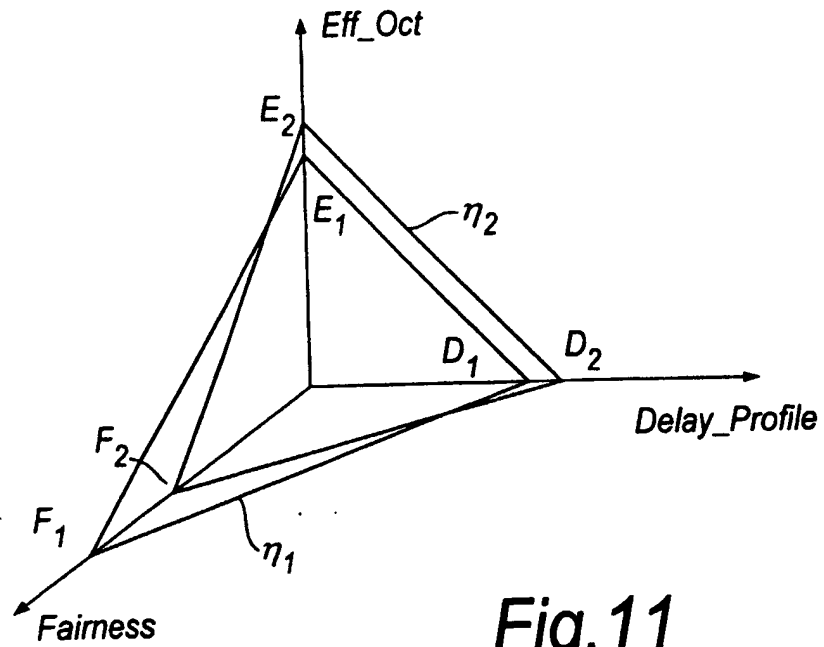


Fig.9

9/23



10/23



11/23

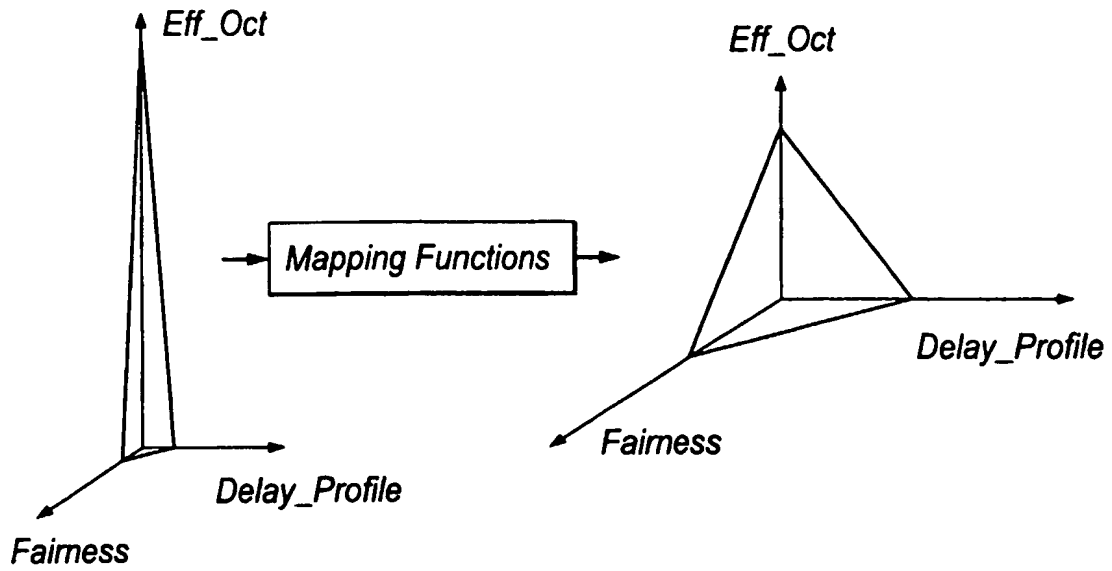


Fig.13

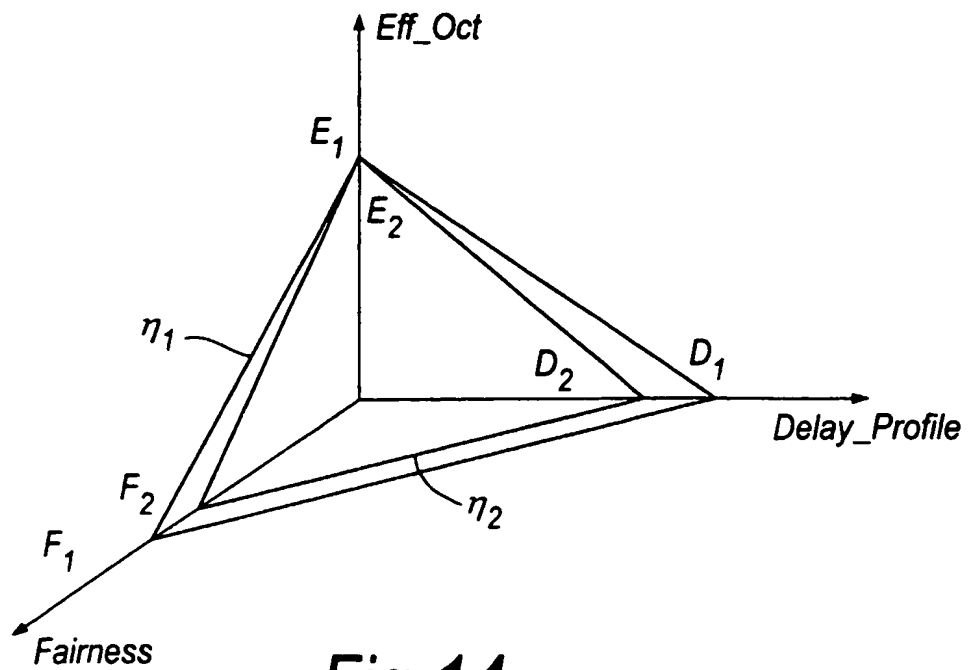


Fig.14

12/23

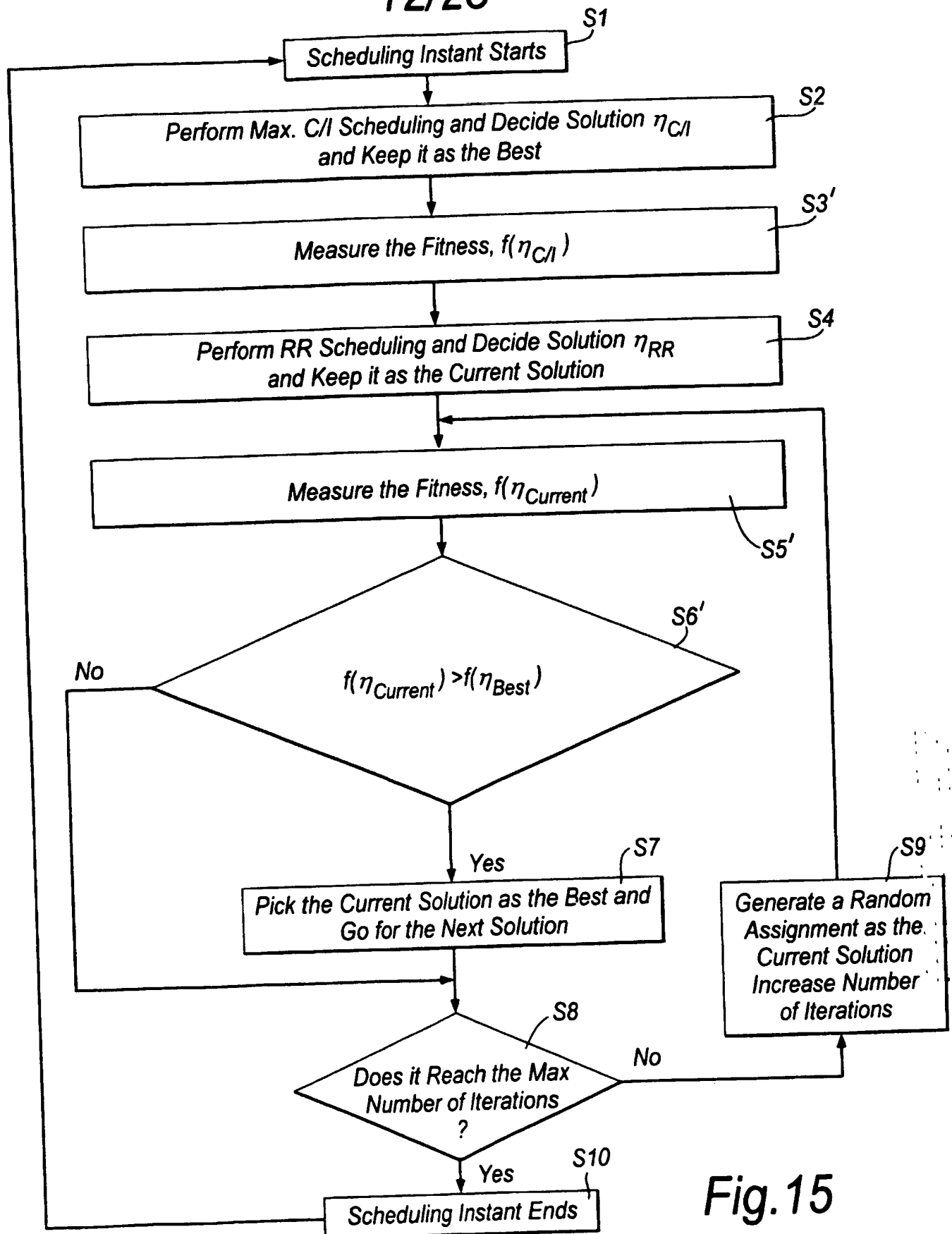


Fig.15

13/23

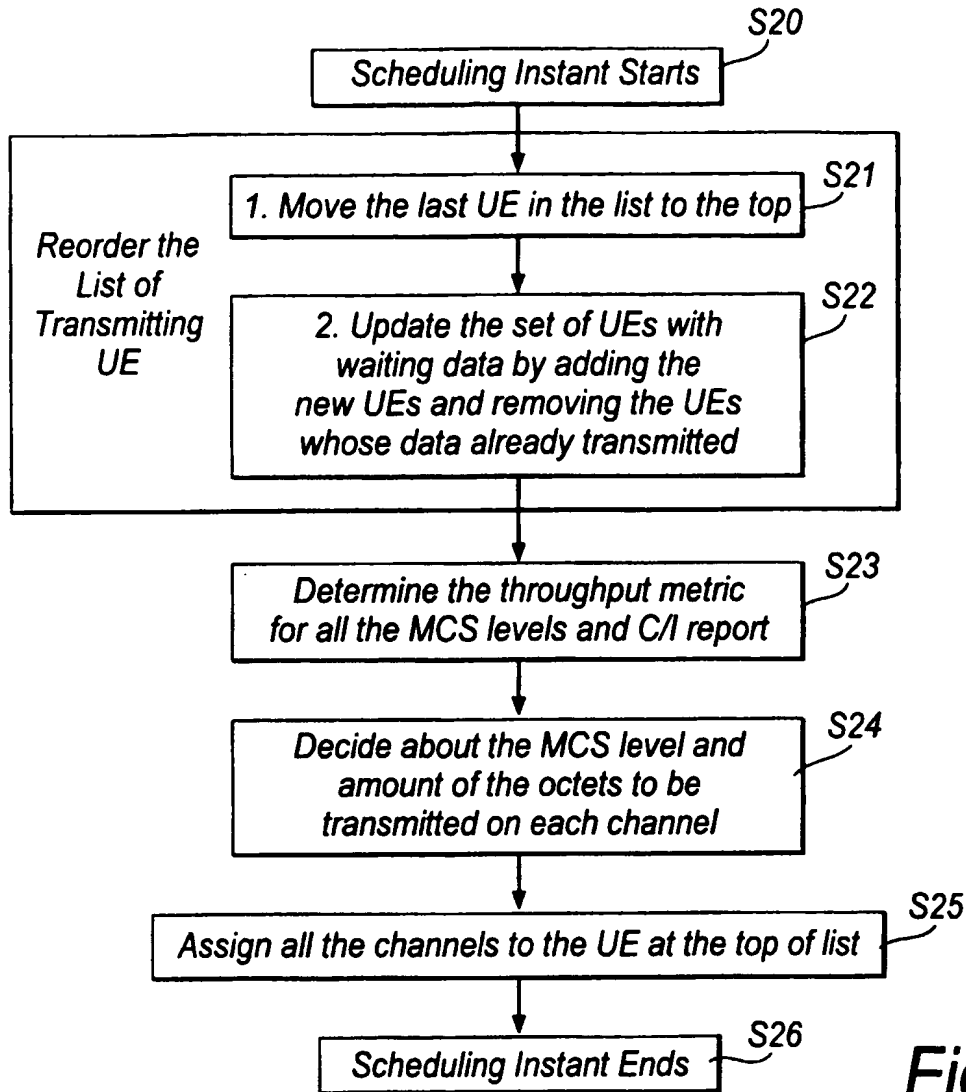


Fig. 16

UE 11
UE 5
UE 32
UE 1
UE 49
UE 8
UE 23
UE 34

Fig. 17(a)

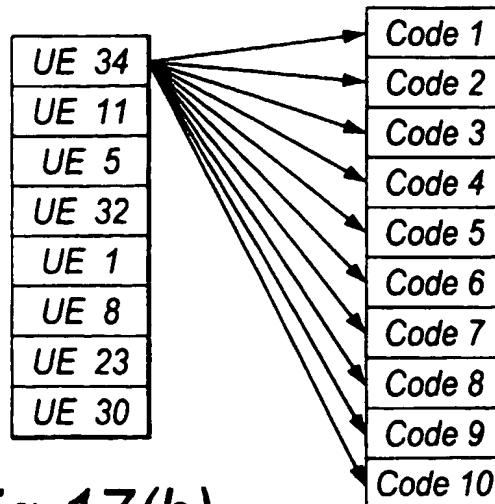


Fig. 17(b)

14/23

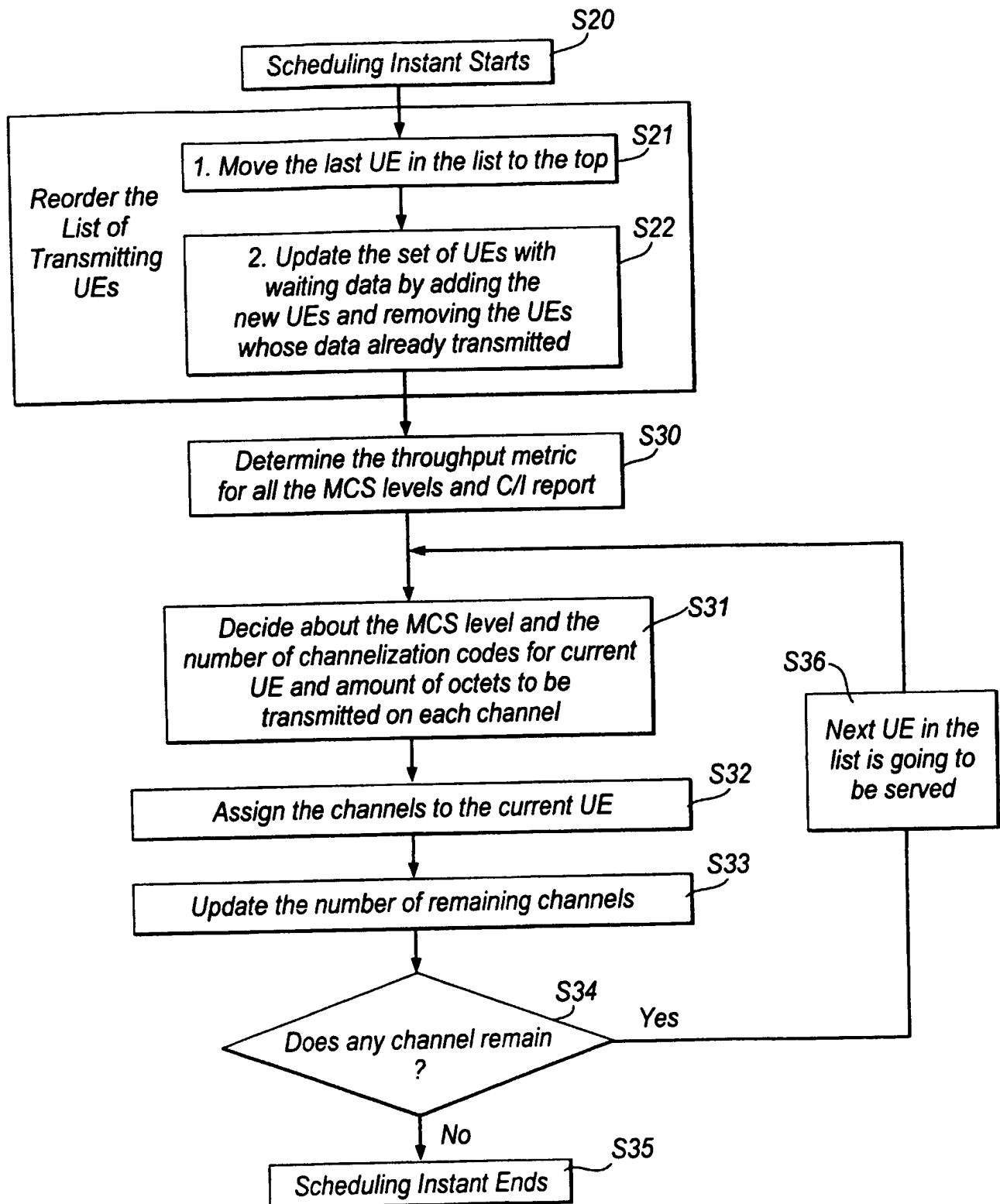


Fig.18

15/23

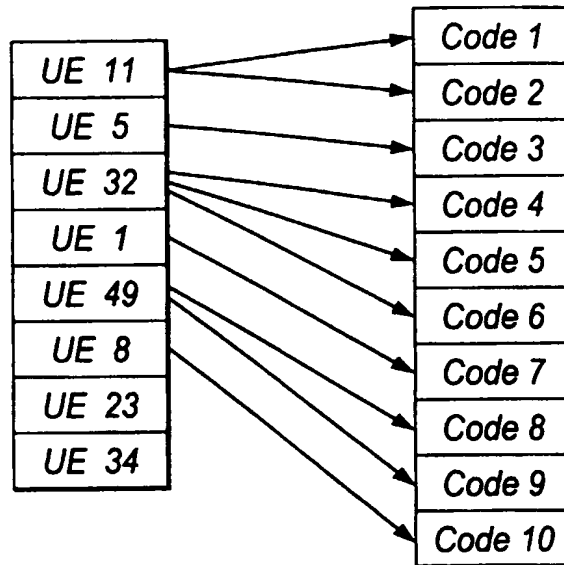


Fig.19(a)

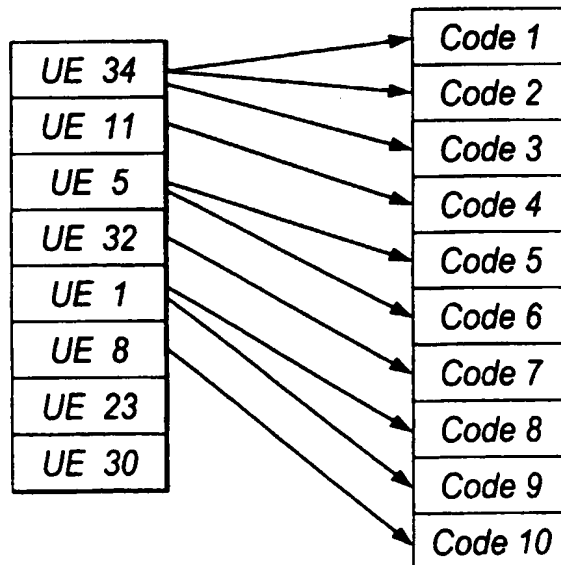


Fig.19(b)

16/23

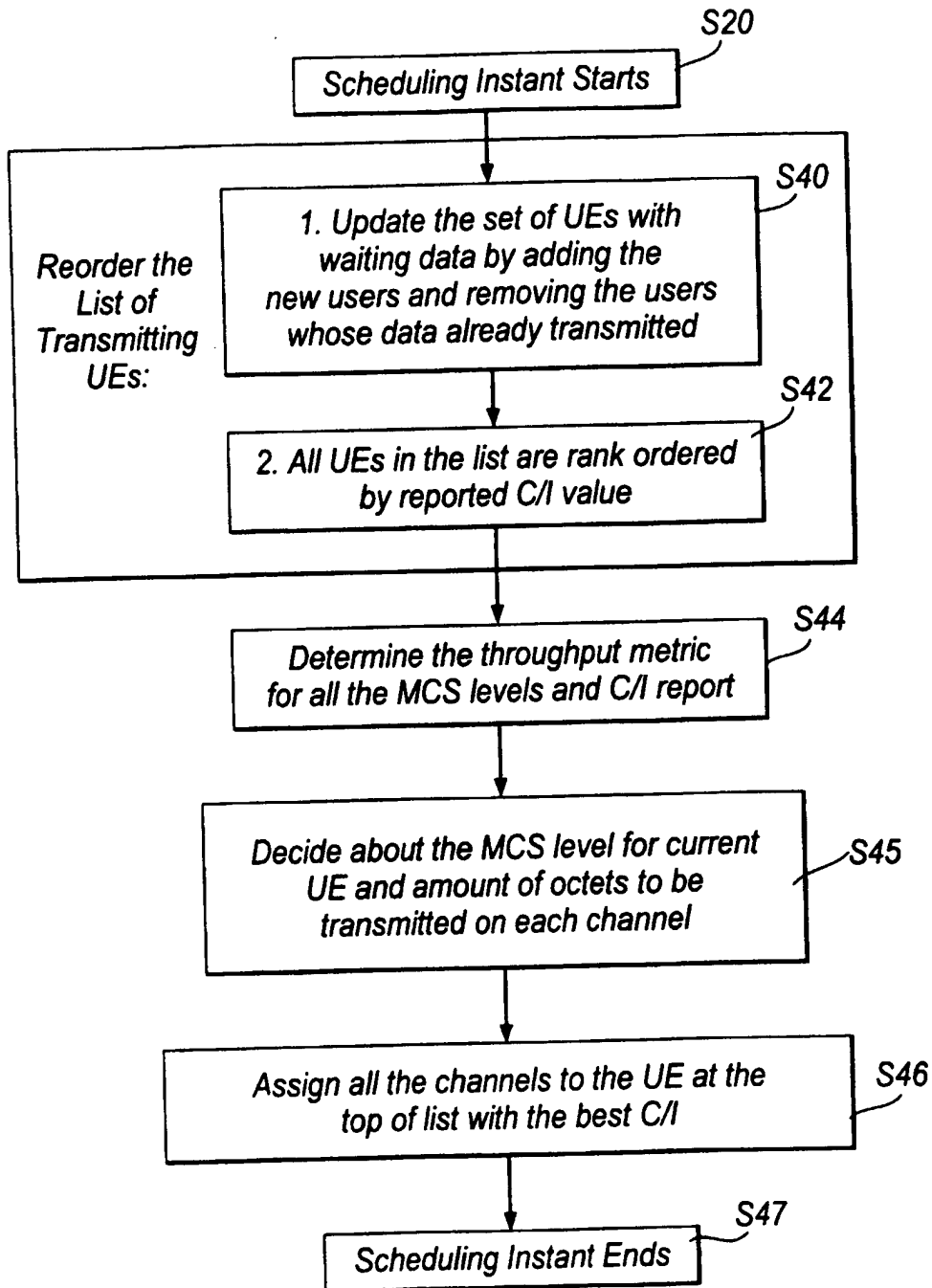


Fig.20

17/23

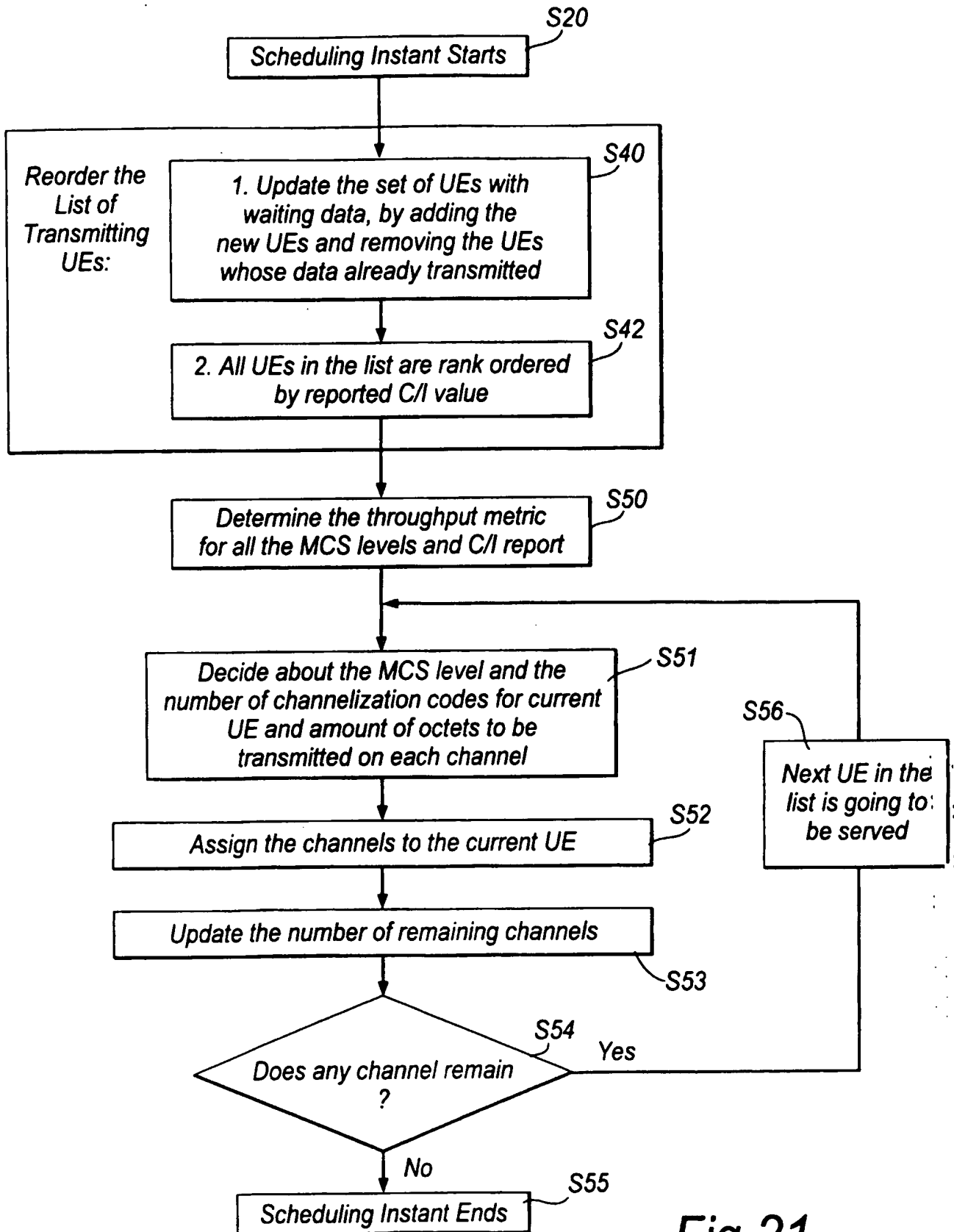


Fig.21

18/23

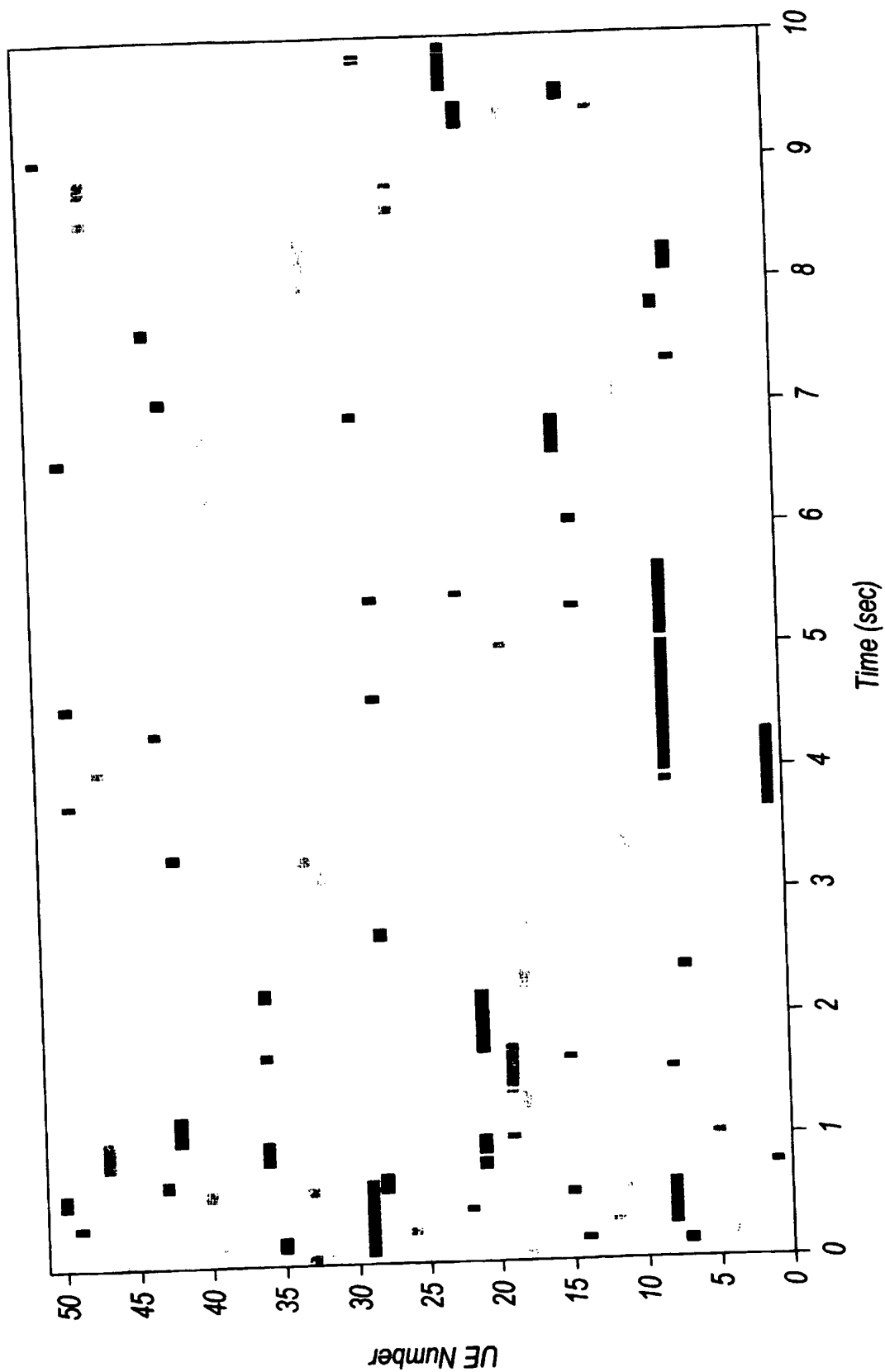
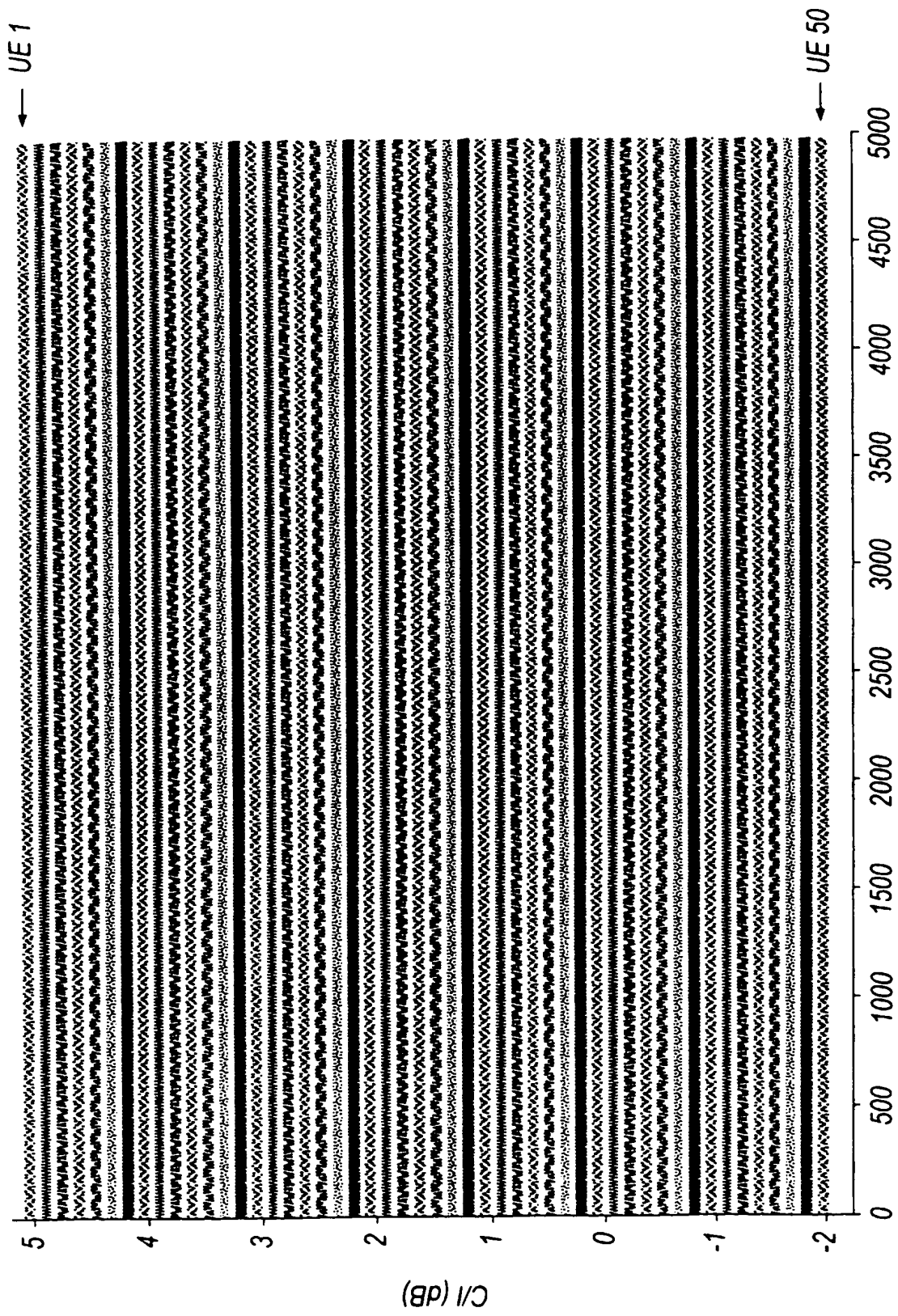


Fig.22



TTI Number

Fig.23

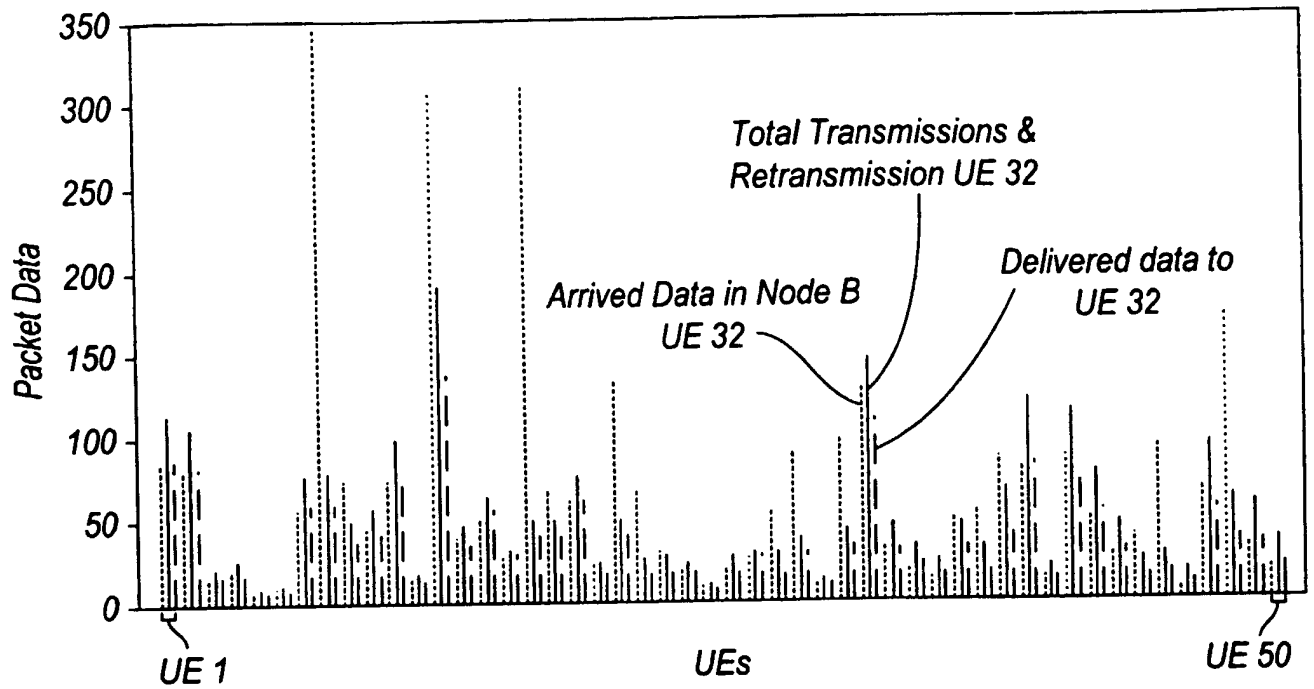


Fig.24

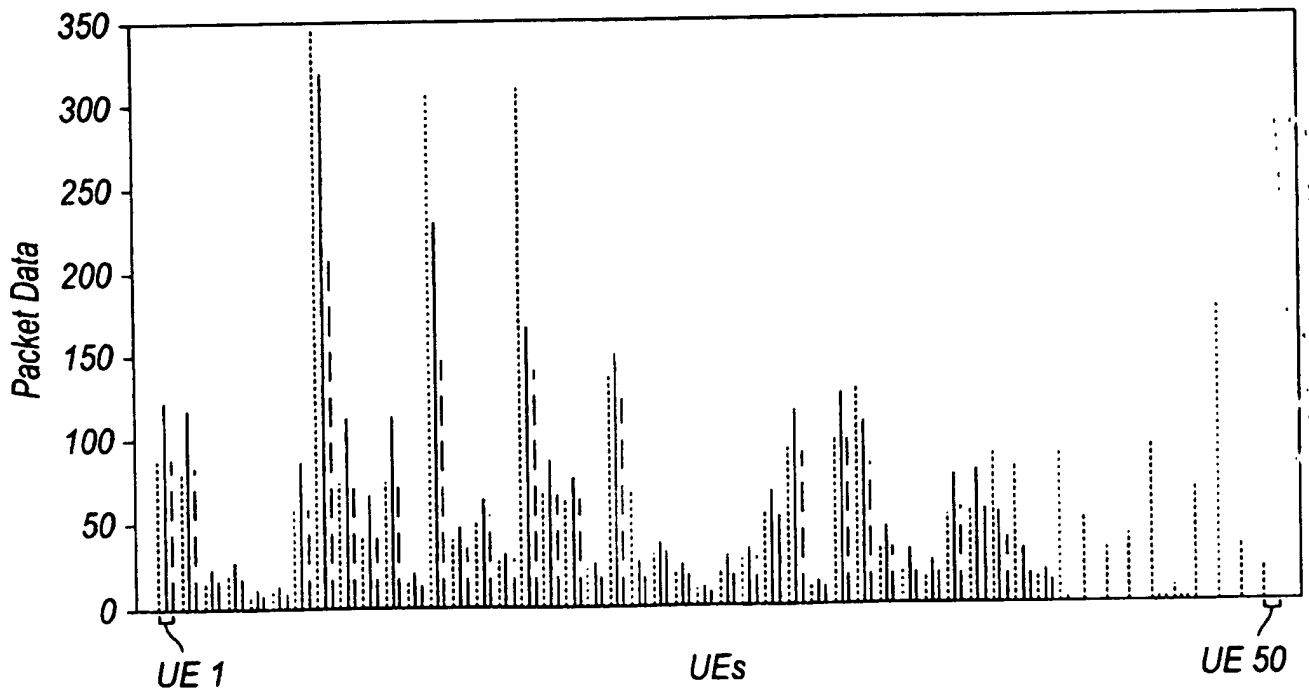


Fig.25

21/23

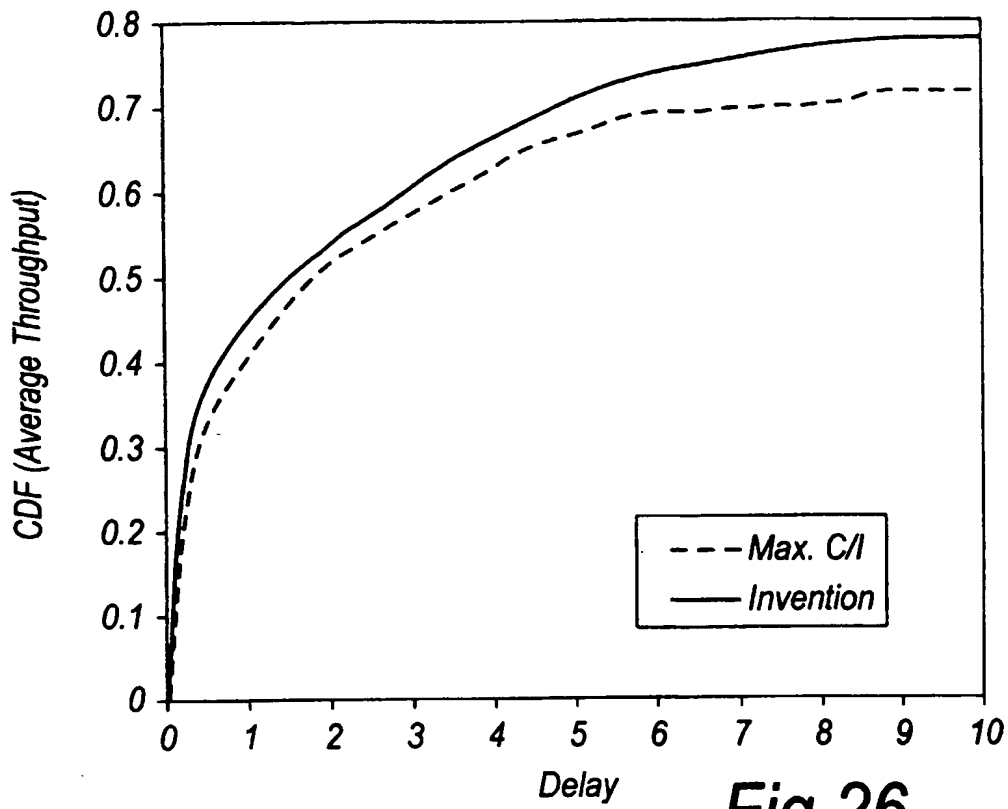


Fig.26

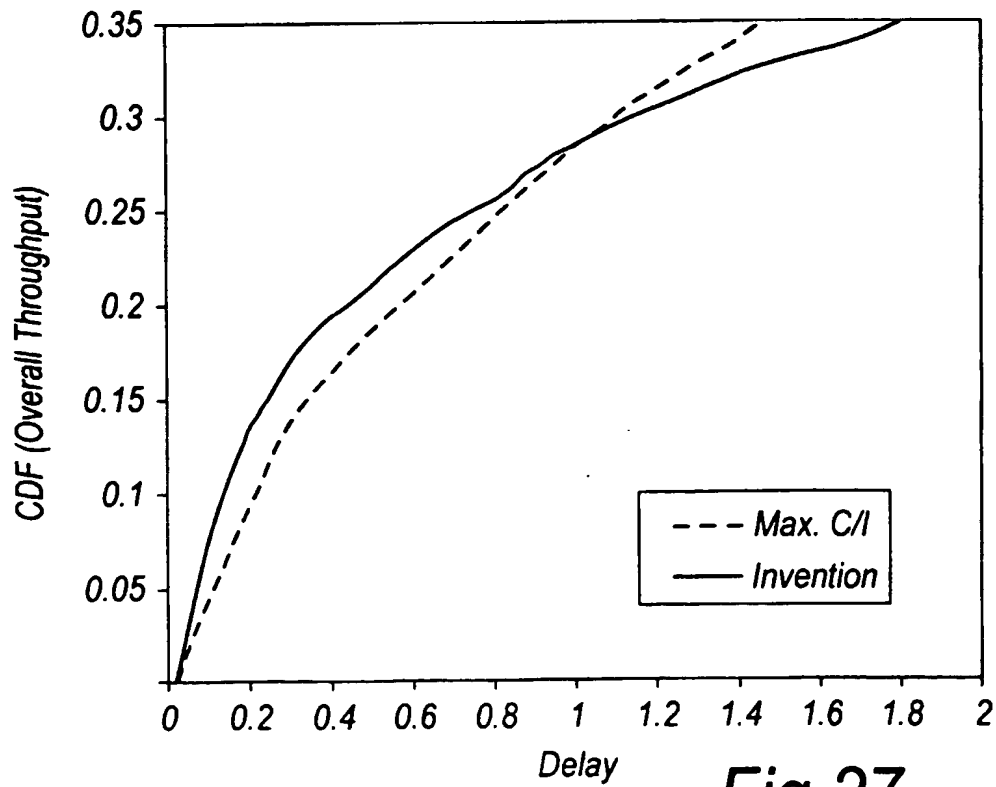


Fig.27

22/23

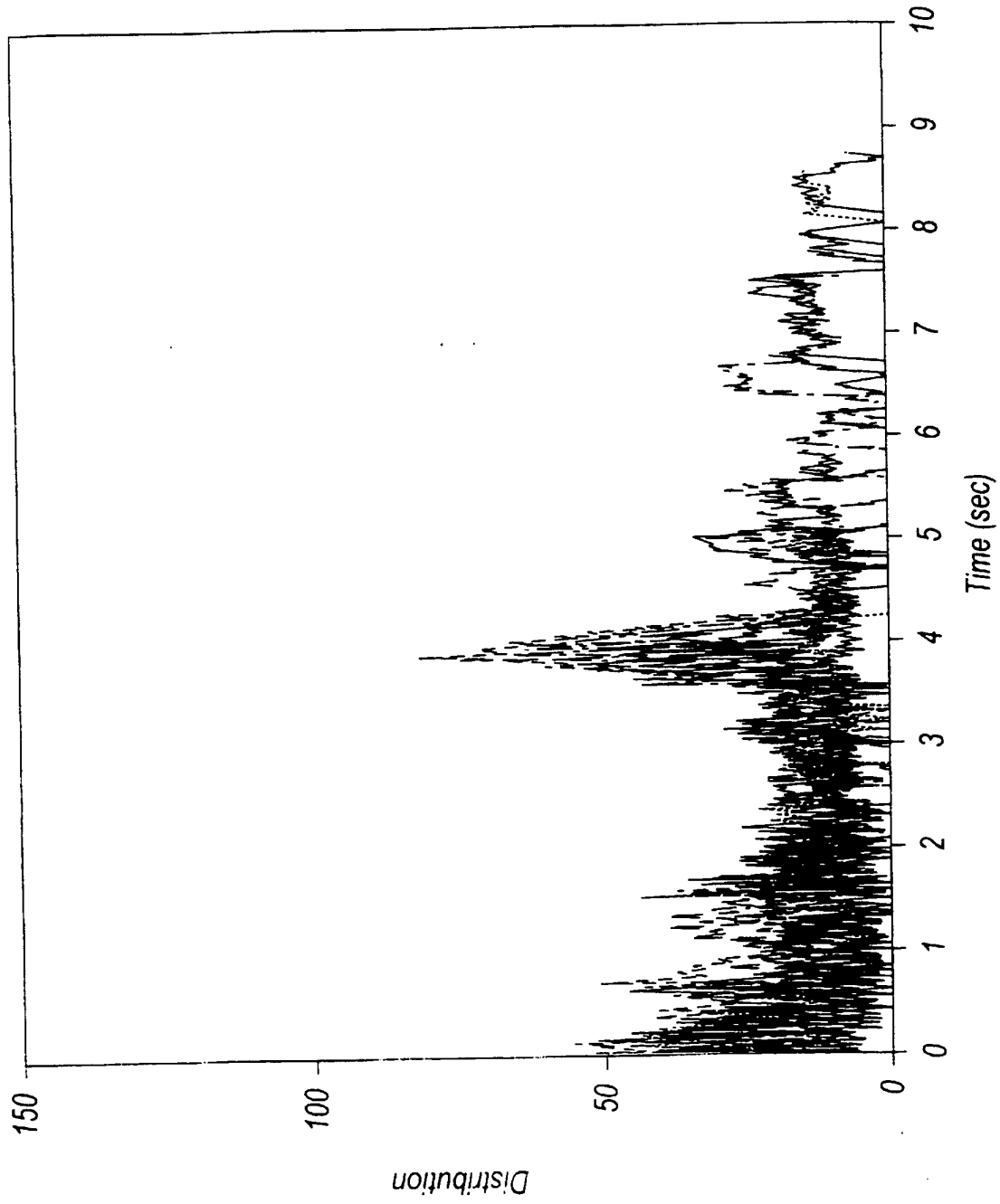


Fig.28

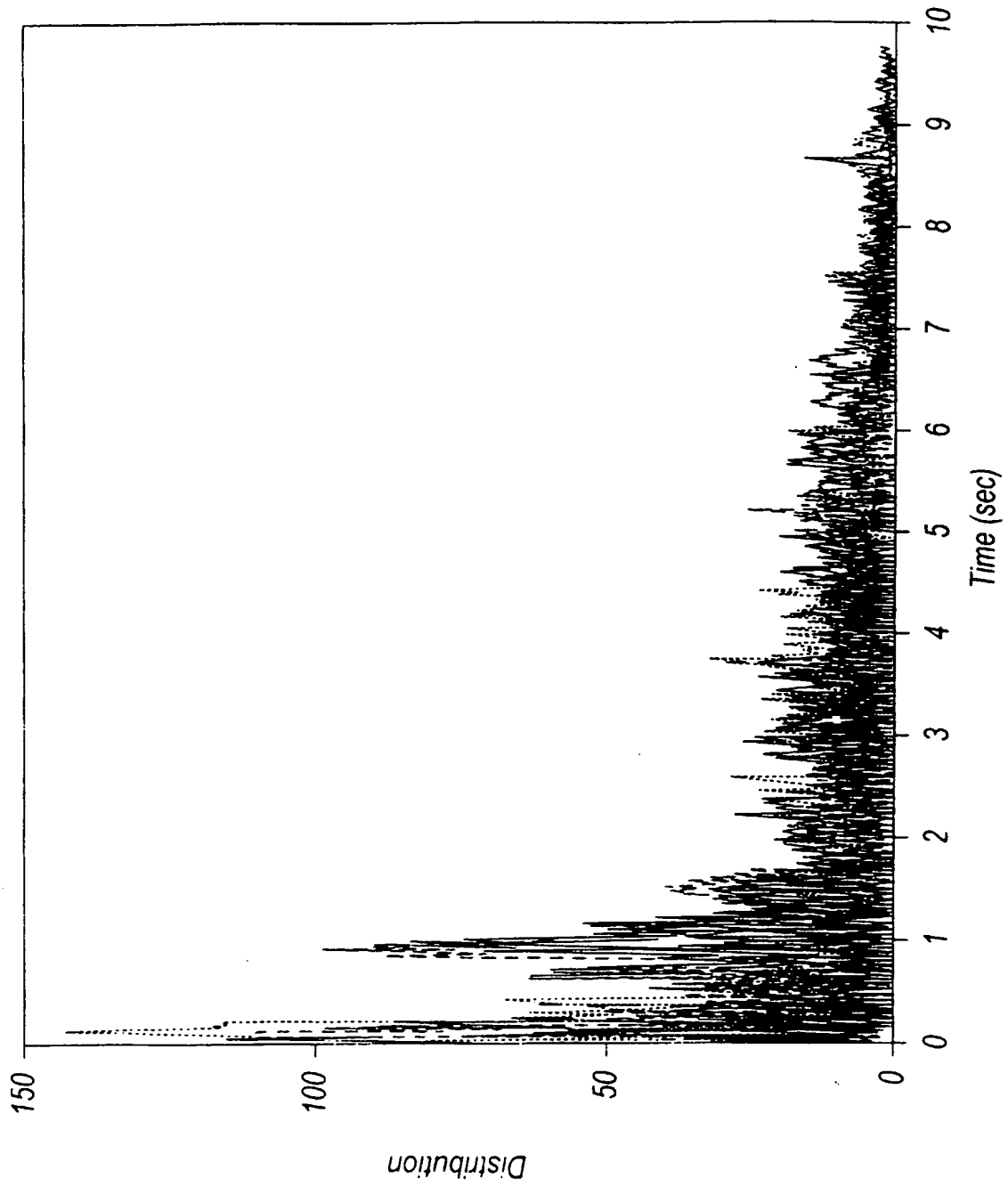


Fig. 29

PACKET SCHEDULING

The present invention relates to packet scheduling methods and apparatus for use, for example, in wireless communication systems.

Fig. 1 shows parts of a wireless communication system 1. The system includes a plurality of base stations 2, only one of which is shown in Fig. 1. The base station 2 serves a cell in which a plurality of individual users may be located. Each user has an individual user equipment (UE). Only the user equipments UE1, UE2 and UE50 are shown in Fig. 1. Each UE is, for example, a portable terminal (handset) or portable computer.

As is well known, in a code-division multiple access (CDMA) system the signals transmitted to different UEs from the base station (also known as "node B") are distinguished by using different channelisation codes. In so-called third generation wireless communication systems a high speed downlink packet access (HSDPA) technique has been proposed for transmitting data in the downlink direction (from the base station to the UEs). In this technique a plurality of channels are available for transmitting the data. These channels have different channelisation codes. For example, there may be ten different channels C1 to C10 available for HSDPA in a given cell or sector of a cell. In HSDPA, downlink transmissions are divided up into a series of transmission time intervals (TTI), and a packet of data is transmitted on each different available channel to a selected UE. A new choice of which UE is served by which channel can be made in each TTI.

Fig. 2 shows an example of the operation of the HSDPA technique over a series of transmission time intervals TTI1 to TTI9. As shown in Fig. 2, in TTI1 it

(
is determined that two packets will be sent to UE50,
four packets will be sent to UE11 and four packets will
be sent to UE2. Accordingly, two channels are
allocated to UE50 and four channels each are allocated
5 to UE11 and UE2. Thus, as shown in Fig. 1, UE50 is
allocated channels C1 and C2, UE11 is allocated
channels C3 to C6, and UE2 is allocated channels C7 to
C10.

In the next transmission time interval TTI2 a new
10 user equipment UE1 is sent one packet, and the
remaining UEs specified in TTI1 continue to receive
packets.

Thus, effectively the HSDPA system employs a
number of parallel shared channels to transmit data in
15 packet form from the base station to the different UEs.
This system is expected to be used, for example, to
support world wide web (WWW) browsing.

In order to decide which UE should be served on
which channel in each TTI a packet scheduling technique
20 is employed. Conventionally, two basic types of
scheduling technique have been considered for use in
HSDPA: a round-robin (RR) scheduling technique and a
maximum carrier-to-interference ratio (max C/I)
technique.

25 The basic round-robin technique first compiles a
list of the UEs which currently have data waiting at
the transmitter (base station) for transmission. For
each TTI the last UE in the list will have the highest
priority for the next TTI. Accordingly, the UEs are
30 serviced in a round robin fashion. In the simplest
round-robin scheduling technique, it is assumed that
the UE with the highest priority takes all of the
channels. However, a packet-weighted round-robin
technique is also known. This allocates the available
35 channels to a group of users in each TTI based on the
relative amounts of data for the different UEs. In

this packet-weighted technique UEs which have more data waiting for transmission are allocated more channels.

The round-robin scheduling techniques emphasise fairness amongst the competing UEs in terms of radio resource allocation. However, they tend to provide relatively poor total throughput of data.

The max C/I scheduling technique is similar to the round-robin scheduling technique except that the list of UEs having waiting data is sorted in each TTI based on a carrier-to-interference ratio (C/I) reported by each UE. The C/I is a measure of the quality of the channel. By sorting the list of UEs based on C/I, UEs which have a better channel quality are given a higher chance to be selected. In the simplest version of the technique, all of the channels are allocated to the UE with waiting data that has the highest C/I. A packet-weighted variant is also possible, in which instead of selecting a single UE having the highest C/I, a group of UEs with the highest C/I values is selected, and the available channels are divided up amongst the group of UEs based on the relative amounts of data which those UEs have waiting for transmission.

The max C/I scheduling technique tends to maximise the total throughput of data but this is at the expense of fairness. It can be seen that UEs which report poor C/I values, for example because they are far from the base station or because there are many other interfering UEs in the vicinity, will only very rarely be selected. Thus, these UEs are likely to suffer from unacceptably long delays in receiving packets.

It is therefore desirable to provide a packet scheduling technique in which no one performance aspect (e.g. throughput or fairness) dominates above all others.

It is also desirable to provide a scheduling technique in which a balance between different

performance aspects (such as throughput, delay and fairness) can be selected by an operator of the system.

It is also desirable to provide a method of determining a unified measure of fitness of a candidate scheduling solution which can enable different candidate scheduling solutions for a particular scheduling instant (transmission time interval) to be compared and a best one of the candidate solutions to be selected.

According to a first aspect of the present invention there is provided a packet scheduling method, for scheduling packets of data for transmission from a transmitter via at least one channel to a plurality of receivers, which method comprises: generating a plurality of candidate scheduling solutions, at least one said candidate solution being generated randomly, and each candidate solution specifying at least the receiver(s) to which packets are to be transmitted in a scheduling instant under consideration; comparing the generated candidate solutions and, based on the comparison results, selecting a best one of the candidate solutions to use to transmit packets in said scheduling instant.

In such a method, a scheduling solution that is superior to that achieved by conventional scheduling methods such as round-robin or max C/I can be generated without the use of significant processing resources.

In one embodiment at least one of the candidate solutions is generated using a deterministic scheduling method. This can enable the deterministic solution(s) to be benchmarks against which the random solutions are compared.

The candidate solutions may include a round-robin scheduling solution generated by a round-robin scheduling method and/or a maximum carrier-to-

(
interference ratio solution generated by a maximum carrier-to-interference ratio scheduling method.

In many practical situations, for example in a wireless communication system, a plurality of channels
5 will be available for transmitting packets from the transmitter to the receivers in each scheduling instant. In this case, each candidate solution preferably further specifies one or more parameters applicable to each available channel in the scheduling
10 instant under consideration.

For example each candidate solution may specify how the specified receiver(s) is (are) to be allocated to the available channels in the scheduling instant under consideration.

15 Alternatively, or in addition, each candidate solution may specify a modulation and/or coding scheme (e.g. MCS level) to be applied to each available channel in the scheduling instant under consideration. This is advantageous if the communication system uses
20 an adaptive modulation scheme to increase the throughput when channel conditions permit.

Alternatively, or in addition, each candidate solution may specify a transmission power for each available channel in the scheduling instant under
25 consideration.

Alternatively, or in addition, each candidate solution may specify an amount of data to be transmitted via each available channel in the scheduling instant under consideration.

30 The method may be carried out over a series of iterations. In this case, in one embodiment in each iteration a new candidate solution is generated and compared with a candidate solution selected in a previous iteration as the best solution. The new
35 candidate solution is selected as a new best solution if, based on the comparison result, it is judged to be

as good as or better than the previously-selected best solution. This reduces memory requirements it avoids the need to store all the candidate solutions: only one best solution is stored at any time.

5 When the plurality of candidate solutions include a round-robin solution determined using a round-robin scheduling method, and a maximum carrier-to-interference ratio solution determined using a maximum carrier-to-interference ratio scheduling method, it is
10 preferable for the maximum carrier-to-interference solution to be generated before an iteration in which the round-robin solution is generated, at least in the case in which throughput is a significant performance aspect taken into account in the comparison. This
15 effectively makes the max C/I solution the dominant solution which can only be replaced by the RR solution if the RR solution has as good a throughput as the max C/I solution (as well as better fairness).

For example, the maximum carrier-to-interference ratio
20 solution may be generated in an initialisation step carried out before the first iteration of the series of iterations, that maximum carrier-to-interference ratio solution being set as the initial best solution against which the new candidate solution generated in the first
25 iteration is compared. The round-robin solution may then be the new candidate solution generated in the first iteration, and the new candidate solutions generated in the second and each subsequent iteration of the series may be randomly-generated solutions.

30 The comparison of the candidate solutions can be carried out in any suitable way. Preferably, however, for each candidate solution, at least one fitness measure is determined, the or each fitness measure being a measure of performance of the candidate
35 solution concerned, and the best solution is identified

based on a comparison of the fitness measure(s) of the generated candidate solutions.

Preferably, a set of two or more such individual fitness measures is determined for each candidate solution, which fitness measures relate to different respective performance aspects. This can enable the method to balance two or more different performance aspects such as throughput and fairness. Conventional scheduling methods either explicitly or implicitly concentrate exclusively on one performance aspect and ignore other aspects.

In one embodiment, for each performance aspect, the corresponding individual fitness measures for the different candidate solutions are compared to select the best solution. This kind of simple logical comparison can be carried out quickly without requiring significant computing resources. Also, there is no need to normalise or otherwise adjust the individual fitness measures to ensure compatibility between the measures for different performance aspects.

The best solution in this case may simply be the solution which has the best fitness measure in every one of the performance aspects. More sophisticated comparisons are possible, however, if desired.

In another embodiment, the two or more individual fitness measures in the set for each candidate solution are combined together using a fitness function to produce a unified measure of fitness for the candidate solution concerned, and the respective unified measures of fitness for the candidate solutions are compared to select the best solution. This enables a simple comparison of unified measures to be used to select the best solution whilst preserving the ability to take into account a balance of performance aspects.

The unified measure of fitness may be produced by forming a product of the individual fitness measures in

(
the set. This is simple but effective. However, when forming the unified measure in this way it may be desirable to adopt one or more of the following techniques to enhance the result.

5 Firstly, at least one the individual fitness measure of the set may be adjusted to reduce a mismatch between it and another one of the individual fitness measures of the set.

10 For example, for one or more the individual fitness measures of the set, a mapping function may be applied to the fitness measure so as to reduce a mismatch between it and another one of the individual fitness measures of the set.

15 Alternatively, or in addition, one or more of the individual fitness measures of the set may be normalised.

20 Secondly, one or more of the individual fitness measures of the set may be weighted when the individual fitness measures are combined together to produce the unified measure of fitness. This enables a desired balance between different performance aspects to be achieved. The weights may be adjustable, if required, to provide a tuning facility and enable operators to choose individually-preferred balances.

25 The performance aspects which can be taken into account are not limited either when carrying out a logical comparison of individual fitness measures or when forming a unified fitness measure. However, the fitness measure may be one or more of the following:

30 a measure of total data throughput to the specified receiver(s) for the candidate solution concerned;

35 a measure of an amount of data at the transmitter waiting for transmission to the specified receiver(s) for the candidate solution concerned;

(
a measure of delay(s) in transmitting data to the specified receiver(s) in the candidate solution concerned; and

5 a measure of how the candidate solution affects a fairness of the packet scheduling amongst active ones of the plurality of receivers.

To cope with different service types (e.g. video data, WWW download data etc.) without undue complexity, the or each delay is preferably calculated relative to
10 a preselected delay parameter applicable to a type of data to be transmitted to the receiver concerned. For example, the delay parameter applicable to a higher-priority type of data may be lower than the delay parameter applicable to a lower-priority type of data.

15 A measure of total throughput may be calculated by estimating a total amount of data which will be transmitted successfully to the specified receiver(s) of the candidate solution in the scheduling instant under consideration.

20 A measure of waiting data may be a ratio of a current amount of data waiting for transmission at the transmitter to each specified receiver to a total the amount of data which has arrived at the transmitter for transmission to the receiver concerned.

25 A measure of delay may be calculated based on a sum of respective delay times for transmissions to the specified receiver(s). Each delay time may be dependent upon a difference between a current time and a time of arrival at the transmitter of the earliest
30 data currently waiting for transmission to the specified receiver.

A fairness measure may be based on a variance or average of respective estimated data throughputs to all active ones of the receivers if the candidate solution
35 is used.

A measure of quality of service may be based on the number of packets delivered to each receiver within a desired threshold of time. The threshold may be different for different service types.

5 The type of transmitter and receiver in the invention is not limited but in one embodiment the transmission is a wireless transmission and the transmitter is part of a base station of a wireless communication system, and each receiver is part of a
10 user equipment of that system. A method and/or apparatus embodying the invention are also applicable when the transmitter is part of the user equipment. A "packet" can be of any suitable length.

 According to a second aspect of the present
15 invention there is provided packet scheduling apparatus, for scheduling packets of data for transmission from a transmitter to a plurality of receivers via at least one channel, which apparatus comprises: solution generating means for generating a
20 plurality of candidate scheduling solutions, at least one said candidate solution being generated randomly, and each said candidate solution specifying at least the receiver(s) to which packets are to be transmitted in a scheduling instant under consideration; best
25 solution selecting means for comparing the generated candidate solutions and selecting, based on the comparison results, a best one of the candidate solutions and to use for transmitting packets in the scheduling instant.

30 According to a third aspect of the present invention there is provided a transmitter comprising: packet scheduling apparatus embodying the aforesaid second aspect of the invention; and transmitting means connected operatively to said packet scheduling
35 apparatus and operable, when scheduling instant occurs, to cause packets to be transmitted to the receiver(s)

specified by the best solution selected for that instant by the packet scheduling apparatus.

Another aspect of the present invention can provide a packet scheduling method, for scheduling
5 packets of data for transmission from a transmitter via at least one channel to a plurality of receivers, which method comprises: generating a plurality of candidate scheduling solutions, each candidate solution
10 specifying at least the receiver(s) to which packets are to be transmitted in a scheduling instant under consideration; producing a measure of fitness for each candidate solution, the fitness measure being a measure of performance of the candidate solution in at least
15 one of the following aspects: delay, fairness, and quality of service; and employing the measures of fitness produced for different candidate solutions to make a selection amongst the generated candidate solutions.

Yet another aspect of the present invention can
20 provide a packet scheduling method, for scheduling packets of data for transmission from a transmitter via at least one channel to a plurality of receivers, which method comprises: generating a plurality of candidate scheduling solutions, each candidate solution
25 specifying at least the receiver(s) to which packets are to be transmitted in a scheduling instant under consideration; producing a measure of fitness for each candidate solution, the fitness measure being a measure of performance of the candidate solution in at least
30 two different aspects; and employing the respective measures of fitness produced for different candidate solutions to make a selection amongst the generated candidate solutions. The different performance aspects may include two or more of the following: throughput,
35 delay, fairness, and quality of service.

(
Preferably, in producing the fitness measure a weighting of at least one the performance aspect is adjustable relative to that of another the performance aspect.

5 Each performance aspect has an individually-adjustable weighting.

The fitness measures are preferably employed to select a best one of the candidate solutions to use to transmit packets to the receivers in the scheduling
10 instant under consideration.

At least one of the candidate solutions may be generated using a genetic algorithm and the fitness measures are employed in a selection process of the genetic algorithm. Some ways of applying a genetic
15 algorithm to scheduling are described in our co-pending United Kingdom patent application no. [agent's ref. HL82318], the entire content of which is incorporated herein by reference, (48 0216245.1).

Reference will now be made, by way of example, to
20 the accompanying drawings, in which:

Fig. 1 discussed hereinbefore, shows parts of a wireless communication system employing a HSDPA technique for downlink transmissions;

Fig. 2 shows an example of the operation of the
25 HSDPA technique in the Fig. 1 system;

Fig. 3 shows a block diagram of packet scheduling apparatus embodying the present invention;

Fig. 4 is a block diagram showing an example of the constitution of one of the parts of the Fig. 3
30 apparatus;

Fig. 5 is a block diagram showing an example of the constitution of another of the parts of the Fig. 3 apparatus;

Fig. 6 is a graph illustrating a example variation
35 in a carrier-to-interference ratio of a downlink channel over a series of transmission time intervals

(
for four different UEs in a wireless communication system;

Fig. 7 is a graph for use in explaining an adaptive modulation and coding technique suitable for use in embodiments of the invention;

Fig. 8 is a schematic diagram for use in explaining an automatic repeat request process suitable for use in embodiments of the present invention;

Fig. 9 is a flowchart illustrating operation of a first embodiment of the present invention;

Figs. 10 to 14 are schematic views for use in explaining a unified fitness measure used in a second embodiment of the present invention;

Fig. 15 is a flowchart for use in explaining operation of the second embodiment of the present invention;

Fig. 16 is a flowchart for use in explaining a scheduling process performed by a generator part shown in Fig. 4;

Figs. 17(a) and 17(b) are schematic views showing an example of the Fig. 16 process;

Fig. 18 is a flowchart for use in explaining a variation of the Fig. 16 process;

Figs. 19(a) and 19(b) are schematic diagrams showing an example of the operation of the Fig. 18 process;

Fig. 20 is a flowchart for use in explaining a scheduling process performed by another generator part shown in Fig. 4;

Fig. 21 is a flowchart for use in explaining a variation of the Fig. 20 process;

Fig. 22 is a diagram illustrating an example of packet transmission activity in an embodiment of the present invention;

Fig. 23 is a schematic diagram for use in explaining carrier-to-interference ratio differences applied to different UEs in a simulation process;

Fig. 24 is a histogram illustrating packet transmission activity for different UEs in a simulation carried out according to the Fig. 23 simulation process for an embodiment of the present invention;

Fig. 25 is a histogram corresponding to Fig. 24 but for a conventional scheduler;

Fig. 26 is a graph illustrating a cumulative density function of an average throughput of data versus delay characteristic for an embodiment of the present invention and for the conventional scheduler or Fig. 25;

Fig. 27 is a graph illustrating a cumulative density function of overall throughput of data versus delay characteristic for an embodiment of the present invention and for the conventional scheduler of Fig. 25;

Fig. 28 is a histogram illustrating a distribution of delivered packets over time for a plurality of different UEs in the conventional scheduler; and

Fig. 29 is a histogram corresponding to Fig. 28 but for an embodiment of the present invention.

Fig. 3 shows a block diagram of packet scheduling apparatus 10 embodying the present invention. The apparatus 10 is used to schedule packets of data for transmission from a transmitter to a plurality of receivers via at least one channel. The transmitter is, for example, a base station (node B) in a wireless communication system. The plurality of receivers in this case are different UEs served by the base station.

The apparatus 10 comprises a candidate solution generator unit 12 which generates a plurality (η) of candidate scheduling solutions η_1 to η_n . Each candidate solution η specifies at least the receiver(s) to which

packets are to be transmitted in a scheduling instant (e.g. transmission time interval TTI) under consideration. At least one of the candidate solutions is generated randomly.

5 The apparatus 10 also comprises a best solution selecting unit 14 which compares the different candidate solutions $\eta_1 \dots \eta_n$ generated by the generating unit 12 and selects a best one of the candidate solutions η_{best} based on the comparison results.

10 The apparatus 10 is connected operatively to a transmission unit 16 which, when a scheduling instant for which the apparatus has selected a best solution occurs, causes packets to be transmitted to the receiver(s) specified in the selected best solution
15 η_{best} .

In the usual case, in which a plurality of channels are available for transmitting packets from the transmitter to the receiver(s), each candidate solution η must further specify how the specified
20 receiver(s) is (are) to be allocated to the available channels in the scheduling instant under consideration.

Fig. 4 shows an example of the constitution of the candidate solution generator unit 12 in the Fig. 3 apparatus. In this example, the candidate solution
25 generator unit 12 comprises first, second and third solution generators 122, 124 and 126.

The first generator 122 is a max C/I solution generator 122 which receives carrier-to-interference ratio (C/I) reports from each active receiver and
30 generates a max C/I candidate solution $\eta_{\text{C/I}}$ for the scheduling instant under consideration based on the reported C/I values. Details of the way in which the max C/I solution is generated by the generator 122 will be given later with reference to Figs. 16 to 19.

35 The second generator 124 is a round-robin (RR) solution generator which generates a candidate solution

(
 η_{RR} for the scheduling instant under consideration using a round robin scheduling technique. Details of the way in which the RR solution η_{RR} can be generated by the generator 124 will be given later with reference to
 5 Fig. 20 and 21.

Incidentally, the data destined for each different receiver is buffered in the transmitter prior to transmission, for example in a transmit queue or source queue corresponding to the receiver concerned.
 10 Depending on the type of scheduling technique desired, information FILL regarding the levels of fullness of the source queues for the different receivers may be provided to the first and second generators 122 and 124 so that the solutions $\eta_{C/I}$ and η_{RR} can take account of
 15 these fill levels. This is required, for example, when the solutions are packet-weighted, as described later with reference to Figs. 18 and 20.

The third generator 126 is a random solution generator which generates a random candidate solution
 20 η_{RM} in which the specified receiver(s) is (are) selected randomly or pseudo-randomly. For example, if there are N receivers the third generator 126 generates an integer between 1 and N for each available channel.

Fig. 5 shows an example of the constitution of the
 25 solution selecting unit 14 in the Fig. 3 apparatus. Each candidate solution η_1 to η_n is supplied to a fitness measure producing unit 142 and a selector 146.

For each candidate solution η_1 to η_n the fitness measure producing unit 142 produces at least one
 30 fitness measure F_1 to F_n . The or each fitness measure for a candidate solution is a measure of performance expected of the candidate solution if it is actually chosen for use in the scheduling instant under consideration. As described later in more detail,
 35 preferably the fitness measure or measures should represent the expected performance of the candidate

(
 solution in two or more difference performance aspects
 which it is desired to balance in the scheduling
 process. For example, the fitness measures may take
 into account one or more of the following performance
 5 aspects: throughput (average and total), delay and
 fairness of packet delivery.

The fitness measures F_1 to F_n for the different
 candidate solutions η_1 to η_n are compared by a fitness
 measure comparing unit 144 which determines which one
 10 of the candidate solutions η_1 to η_n has the best fitness
 measure (or best set of fitness measures) and outputs a
 control signal "best" indicating the identified
 solution. The selector 146 receives the control signal
 "best" and selects the solution η_{best} from among the
 15 candidate solutions η_1 to η_n .

Prior to describing preferred embodiments of the
 present invention in more detail, some background
 information regarding the HSDPA system will be
 provided.

20 In the HSDPA system, channel state information
 (CSI) is made available to both the transmitter and the
 receiver, in order to realise a robust communication
 system structure. The HSDPA system is intended to
 increase the transmission rates and throughput, and to
 25 enhance the quality of service (QoS) experienced by
 different users. It transfers most of the functions
 from the base station controller (also known as the
 radio network controller or RNC) to the base
 transceiver station (node B).

30 In addition to employing a number of parallel
 shared channels as described in the introduction with
 reference to Fig. 2, the HSDPA system may also seek to
 achieve high transfer rates using other control
 techniques, as will now be explained.

35 One such control technique, referred to as a
 modulation and coding scheme (MCS), enables the

transmitter to select different modulation and/or coding schemes under different channel conditions.

The C/I value for a channel between the transmitter and a receiver (UE) varies significantly over time. Fig. 6 shows an example of the variation of the C/I values for four different users over a series of 5000 TTIs. This plot was obtained by a simulation. As illustrated, for a given UE the range of C/I values may be as much as from around +12dB to -15dB. The C/I value varies due to shadowing, Rayleigh fading, and change in distribution of the mobile UEs, as well as cellular area specifications including the propagation parameters and speeds of UEs.

Fig. 7 is a graph illustrating a relationship between a data transmission rate and a channel-to-interference ratio (C/I) for four different modulation and coding combinations. The first three combinations are all quadrature amplitude modulation (QAM) schemes which differ from one another in the number (64 or 16) of constellation points used. The fourth combination uses quadrature phase shift keying (QPSK) as its modulation scheme.

Each combination uses coding defined by a coding parameter which, in this example, is expressed as a redundancy rate R . For the first two combinations the redundancy rate R is $3/4$, and for the third and fourth combinations the redundancy rate is $1/2$.

As can be seen from Figure 7, for C/I values lower than around -8dB the fourth modulation and coding combination (QPSK, $R=1/2$) is the only possible option. The characteristic of this combination is plotted with crosses in the figure.

For C/I values in the range from around -8dB to around -2dB, the third combination (16QAM, $R=1/2$) provides the best transmission rate. The

(
characteristic for this combination is illustrated by triangular points in the figure.

For C/I values between around -2dB and +4dB the second combination (16 QAM, $R=3/4$) provides the best
5 transmission rate. The characteristic of this combination is illustrated by square points in the figure.

Finally, for C/I values greater than around +4dB, the first combination (64 QAM, $R=3/4$) provides the best
10 transmission rate. The characteristic of this combination is illustrated by round points in the figure.

The different combinations in Fig. 7 may be referred to as different MCS levels.

15 In the HSDPA system a technique such as adaptive modulation and coding (AMC) is used to adapt the MCS level in accordance with the variations of the channel condition (e.g. C/I value). Each UE produces a measure of the C/I of a downlink channel from the base station,
20 and reports this measure (C/I value) to the base station. The base station then employs the reported C/I values for each UE, as well as information relating to the system limitations and available MCS levels, to identify the most efficient MCS level for the
25 particular UE. Thus, UEs that have better channels or are located in the vicinity of the base station can employ higher levels of MCS and therefore enjoy higher transmission rates. The selection can be carried out, for example, by imposing thresholds (e.g. -8dB, -2dB,
30 +4dB in the example illustrated with reference to Fig. 7) for moving to the next MCS level. Effectively, the result is a classification of the transmission rates based on the channel quality of each UE.

Ideally, each UE reports a C/I value in every TTI
35 and the base station is capable of setting a new MCS level for each available channel in every TTI. This

means that each candidate solution η should preferably include information specifying the MCS level to be applied to each available channel in the TTI under consideration. This MCS level is determined based on
5 the latest reported C/I value for the receiver which, according to the candidate solution η concerned, is to receive a packet in the TTI under consideration.

AMC enables coarse selection of the transmission rate in the HSDPA system. Another control technique
10 which provides for a fine tuning ability of the data rate based on the channel conditions is referred to as a hybrid automatic repeat request (H-ARQ) technique.

Fig. 8 is a schematic diagram for use in explaining how the H-ARQ technique works. In this
15 example, the technique is a so-called stop-and-wait (SAW) version of the technique. The figure shows packet transmissions in a single downlink channel HSPDSCH1 over a series of successive TTIs, TTI1 to TTI9. In TTI2 a first packet is transmitted to UE1.
20 Upon receiving a packet, each UE checks whether the transmission was error-free. If so, the UE sends an acknowledge message ACK back to the base station using an uplink control channel such as the dedicated physical control channel (DPCCH). If there was an
25 error in the transmission of the received packet, the UE sends a non-acknowledge message NACK back to the base station using the uplink channel.

In the example shown in Fig. 8, the first packet transmitted to UE1 in TTI2 fails to be received error-
30 free, and accordingly some time later, in TTI4, UE1 sends the NACK message to the base station. In the H-ARQ technique it is permitted for the next packet destined for a particular UE to be transmitted without waiting for the acknowledge or non-acknowledge message
35 of a packet previously transmitted to the same UE. Thus, none of the transmission timeslots can go idle in

(the case of error-free channels, which gives the ability to schedule UEs freely. System capacity is saved while the overall performance of the system in terms of delivered data is improved.

5 For example, as shown in Fig. 8, before the NACK message for the first packet of UE1 is received by the base station, the base station transmits a second packet to UE1 in TTI4. Thus, this second packet for UE1 is transmitted before the first packet for UE1 is
10 retransmitted in TTI7 in response to the NACK message for the first transmission of the first packet.

In the H-ARQ technique, an erroneously-received packet (failed packet) is subject to a so-called chase combining process. In this process a failed packet is
15 resent by the transmitter and subsequently the receiver "soft" combines (for example using maximal ratio combining) all received copies of the same packet. The final carrier-to-interference ratio (C/I) is determined as the sum of the respective C/Is of the two packets
20 being combined. Thus, the chase combining process improves the C/I of the transmitted packets.

All retransmissions have a higher priority than first transmissions. This means that all retransmission packets are given the opportunity to be
25 transmitted before the first transmissions of new packets.

The HSDPA system specifies the number of retransmissions permitted. If a packet cannot be delivered error-free within the permitted number of
30 retransmissions, it will be dropped. The number of permitted retransmissions is a parameter that has a high impact on the overall system performance. Increasing the number of retransmissions results in a better C/I, which in turn improves the frame error-rate
35 (FER). However, since retransmissions have higher priority than first transmissions, increasing the

permitted number of retransmissions will also increase the delivery delay and have a negative impact on the throughput. Accordingly, in the HSDPA system a trade-off exists between the number of permitted

5 retransmissions, the throughput and the delay of the system. Incidentally, the MCS level for first transmissions and retransmissions may be different.

Embodiments of the present invention seek to optimise the performance of the HSDPA system in terms
10 of measures such as QoS, fairness of packet delivery, total delay, total throughput and average throughput. The optimum channel assignment and packet scheduling policy can be defined as

$$\max_{(\overline{Oct})_k, (\overline{MCS})_k, (\overline{Po})_k, (\overline{UE})_k} f(Th, Fa, 1/De, QoS, Avg_Th) \Big|_{r, (\overline{CTI})_k}, \quad (1)$$

$$k = 1 \dots \alpha$$

15

where α is the number of TTIs, $(\overline{Oct})_k$ is the decision for the number of octets to be transmitted in the k th TTI on each of C channels codes so that

$$\overline{Oct} = \{Oct_1, Oct_2, \dots, Oct_C\}, \quad (2)$$

20

where Oct_i is the number of octets assigned to channel i . Similarly $(\overline{MCS})_k$ defines the decision on the MCS levels for k th TTI for the C channels codes so that

25

$$\overline{MCS} = \{MCS_1, MCS_2, \dots, MCS_C\}, \quad (3)$$

where MCS_i is the MCS level assigned to channel i . $(\overline{UE})_k$ defines the decision on the UE which will be assigned to which channel so that the vector \overline{UE} is

30

$$\overline{UE} = \{UE_1, UE_2, \dots, UE_C\}, \quad (4)$$

where UE_i is the candidate UE to whom the data is to be delivered on channel i . Vector $(\overline{C/I})_k$ is employed to represent the reported C/I values on each channel for
 5 k th TTI

$$\overline{C/I} = \{C/I_1, C/I_2, \dots, C/I_c\}. \quad (5)$$

where C/I_i is the reported C/I value on channel i . We define Th in (1) as the total throughput of the HSDPA
 10 system

$$Th = \frac{\sum_{n=1}^F (Oct_{Received})_{UE_n}}{\sum_{n=1}^F (Oct_{Arrived_Node_B})_{UE_n}}, \quad (6)$$

where F is the number of UEs with active sessions, $(Oct_{Received})_{UE_n}$ is the number of successfully delivered
 15 octets to the n th UE and $(Oct_{Arrived_Node_B})_{UE_n}$ is the number of octets originally delivered to the source queue of n th UE at the transmitter (Node B). The individual throughput of each UE is defined as

$$Th_n = \frac{(Oct_{Received})_{UE_n}}{(Oct_{Arrived_Node_B})_{UE_n}}, \quad n = 1 \dots F, \quad (7)$$

20

and the vector of the individual throughputs as

$$\overline{Th} = \{Th_n\}, \quad n = 1 \dots F, \quad (8)$$

The mean of throughputs or Avg_Th as
 25

$$Avg_Th = mean(\overline{Th}), \quad (9)$$

and Fa or fairness as the variance of \overline{Th}

$$Fa = \text{var}(\overline{Th}), \quad (10)$$

where $\text{var}(\bullet)$ represents the variance operation.

To measure the delivery delay, each octet of packet data may be stamped on its arrival at the Node B. The time at which it is correctly delivered to the UE is recorded. The delivery delay is defined as the difference between the arrival time at the Node B and the successful delivery. This delay may be defined as

$$\text{delay}_i = (\text{delivery_time}_i)_{UE} - (\text{arrival_time}_i)_{Node\ B}, \quad (11)$$

10

In (1) the total delay De for HSDPA system may be defined as

$$De = \sum_{i=1}^{\theta} \text{delay}_i, \quad (12)$$

15 where θ is the total number of the received octets for all UEs.

A QoS parameter in (1) may be defined as the ratio of the successfully delivered data in a defined threshold of time or γ to the total number of transmitted octets. This defined threshold of time γ is preferably different for different types of data. For example, voice or video data will have a lower threshold than an e-mail message or WWW download. The delays calculated in (11) are preferably also expressed relative to the threshold γ applicable to the data type.

25 Finally $(\overline{Po})_k$ is defined as the power allocated to each channel so that

$$\overline{Po} = \{Po_1, Po_2, \dots, Po_C\}, \quad (13)$$

30

where Po_i is the power allocated to channel i .

(
 $f(\bullet)$ is a function (fitness function) that combines all the QoS related parameters to a unique value, as described later in more detail.

A first embodiment of the present invention will now be described in more detail with reference to Figs. 9 and equations (14) to (24).

In this embodiment it is assumed that the allocated powers for the channels are equal, i.e.

$$Po_1 = Po_2 = \dots = Po_C, \quad (14)$$

10

Looking at the optimum scheduler and channel assignment in (1) it can be seen that the optimum assignment requires three vectors $(\overline{Oct})_k$, $(\overline{MCS})_k$ and $(\overline{UE})_k$ with length α . It is preferable to try to construct these vectors as time progresses for each TTI. Thus, in this embodiment each candidate solution is represented as

$$\eta = \{\overline{UE}, \overline{C/I}, \overline{Oct}, \overline{MCS}\}, \quad (15)$$

20 The MCS levels, \overline{MCS} , are decided based on the C/I reported values $\overline{C/I}$. Finally depending on \overline{MCS} levels and the transmission or retransmission states a decision is made on the number of available octets per channel $(\overline{Oct})_k$.

25 Fig. 9 is a flowchart for use in explaining operation of the first embodiment. The processing shown in Fig. 9 is performed for each scheduling instant under consideration. Each scheduling instant corresponds to one transmission time interval (TTI), for example. Of course, the scheduling required for a particular scheduling instant (TTI) must be determined in advance of the time at which that scheduling instant actually occurs, because the selected scheduling

30

solution η_{best} is needed at the start of the TTI concerned.

In a first step S1 the processing for the scheduling instant under consideration is commenced.

- 5 In step S2, the first generator 122 in the candidate solution generating unit 12 is activated to produce a max C/I solution $\eta_{\text{C/I}}$. For reasons that will be apparent later, this solution $\eta_{\text{C/I}}$ is provisionally selected as an initial best solution η_{best} .

- 10 Next, in step S3 a set of individual fitness measures F for the max C/I solution $\eta_{\text{C/I}}$ is produced by the fitness measure producing unit 142 in the solution selecting unit 14.

- 15 In this embodiment the fitness of each solution is evaluated by estimating the contribution of the solution to throughput, delay and fairness.

As the first fitness measure the unit 142 estimates the number of the delivered octets if the current solution η is chosen so that

20

$$Eff_Oct(\eta) = \sum_{i=1}^C \overline{Oct}(i) \cdot (1 - FER(UE(i))) , \quad (16)$$

- where $UE(i)$ is the UE assigned to channel i and $FER(UE(i))$ is the estimated FER for channel i . When the frame error rate in a channel increases the possibility of receiving the error free octets decreases.
- 25

- As the second fitness measure for each solution η , in the current TTI and for the candidate UE s, the unit 142 measures the ratio of octets waiting at the transmitter octets for transmission to the total number of octets arrived in the Node B source queue of that UE so far, so that
- 30

$$\text{Ratio_Waiting_Oct}(\eta) = \frac{\sum_{i=1}^{\omega} \text{Waiting_Oct}(\text{Selected_UE}(i))}{\text{Arrived_Oct}(\text{Selected_UE}(i))} \quad (17)$$

where *Selected_UE* is the set of the selected UEs in solution η and ω is the number of candidate UEs in the current solution so that

$$1 \leq \omega \leq C \quad (18)$$

As the third fitness measure the unit 142 measures the delay experienced by the earliest undelivered octet which has arrived at the transmitter but has not yet been delivered. This measure is defined as

$$\text{Delay_Profile}(\eta) = \frac{1}{\sum_{i=1}^C (N.TTI - \text{Arrival_Time_Earliest}(\text{UE}(i)))} \quad (19)$$

where *N.TTI* represents the time at the current TTI, *Arrival_Time_Earliest(.)* is a vector which includes the arrival time of the earliest undelivered octet in Node B for each UE, and *UE(i)* is the UE assigned to channel *i* for the current solution η .

Finally the fourth fitness measure is defined based on the fairness of the channel assignment and scheduling process for the HSDPA system. The fourth fitness measure is effectively an estimate of the impact of the candidate solution on the fairness of packet scheduling process. For the current TTI, the number of successfully received octets with ACK messages is determined. It is defined as

$$\text{Received} = \{R_i\}_{i=1 \dots F} \quad (20)$$

where F is the number of UEs with active transmission sessions, i is the UE number and R_i is the number of successfully received octets for i th UE. The estimated received octets for the current solution η is defined as

$$\text{Estimated_Received}(\eta) = \{\hat{R}_i\}_{i=1 \dots F} \quad (21)$$

where F is the number of UEs with active transmission sessions, i is the UE number and \hat{R}_i is the number of successfully received octets for i th UE so that for each solution η

$$\hat{R}_{\text{Selected_UE}(c)} = R_{\text{Selected_UE}(c)} + \overline{\text{Oct}}(c), \quad c = 1 \dots C, \quad (22)$$

where c indicates the channel number. Based on the estimated number of received octets, for the current solution η and the current TTI, the expected throughput is estimated for all F active users so that

$$\hat{T}h(\eta) = \left\{ \frac{\hat{R}_i}{\text{Arrived_Oct}(i)} \right\}_{i=1 \dots F}, \quad (23)$$

The expected contribution of the current solution to the fairness of the packet scheduling process is defined as

$$\text{Fairness}(\eta) = \frac{1}{\text{var}(\hat{T}h(\eta))}, \quad (24)$$

After determining these four fitness measures for the max C/I solution $\eta_{C/I}$ in step S3, in step S4 the second generator 124 in the candidate solution generating unit 12 is activated to reduce a round robin solution η_{RR} for the scheduling instant under consideration. This solution η_{RR} is made the current

(solution for comparison against the initially-selected best solution $\eta_{c/1}$.

The following steps S5 to S9 are then carried out iteratively. The number of iterations is, for example,

5 100.

In step S5 the four fitness measures are determined for the current solution. In step S6 a logical comparison is made between, on the one hand, the four fitness measures for the current solution, and, on the other hand, the four fitness measures for the existing best solution. If, and only if, every one of the four fitness measures for the current solution is greater than or equal to the corresponding fitness measure for the existing best solution is the current solution selected as the new best solution in step S7. Otherwise, the current solution is discarded and the existing best solution retained.

In step S8, it is checked whether an iteration count has reached the predetermined number of iterations. If not, in step S9 the third generator in the candidate solution generating unit 12 is activated to produce a random solution η_{RM} which becomes the current solution for the next iteration. The iteration count is incremented and the next iteration begins.

If it is found in step S8 that the predetermined number of iterations has been carried out, processing proceeds to step S9 in which the processing for the scheduling instant under consideration is terminated and the best solution found over the course of the iterations is output as the solution η_{best} . Processing then proceeds back to step S1 for the next scheduling instant under consideration.

In the first embodiment, the candidate solutions are generated and compared using an iterative process (steps S5 to S9). This avoids the requirement to store

(all the candidate solutions and their associated fitness measures, so that this embodiment is efficient in terms of memory usage. It will be appreciated, however, that if memory is not limited, then in a first
 5 stage all the candidate solutions could be generated and the fitness measures for all the candidate solutions could be determined. In a second stage, the stored fitness measures for all the different candidate solutions could be compared, either sequentially or
 10 even in parallel, to identify the best solution.

Also, in the first embodiment it will be observed that the max C/I and round robin solutions $\eta_{C/I}$ and η_{RR} effectively provide the boundaries for the solution in the scheduling instant under consideration.

15 In the logical comparison of step S6 it is not necessary to compare all four fitness measures for the individuals. For example, if throughput is not important the first fitness measure could be omitted from the comparison process. Any desired fitness
 20 measures or combination of fitness measures may be compared in step S6 according to the performance aspects desired to be taken into account.

It will be appreciated that in the first embodiment the comparison of the fitness measures in
 25 step S6 is a simple logical comparison. This is quick and convenient to perform. In particular, no adjustment or normalisation of the fitness measures relative to one another is required, since each fitness measure for the current candidate solution is only
 30 compared with the corresponding fitness measure of the existing best solution.

However, the logical comparison scheme suffers from the disadvantage of being a crude comparison. For example, if the current solution is better than the
 35 existing best solution on three out of the four fitness measures, it will be rejected even if it is only

marginally worse than the best solution on the remaining measure. Under these circumstances, it might actually be better to select the current solution as the new best solution.

5 Of course, the logical comparison could be modified to take account of this kind of situation. For example, if three out of four fitness measures are better for the current solution, the current solution could be selected if the remaining measure is not more
10 than a certain amount worse than the existing best solution. Alternatively, if all the fitness measures are comparable (e.g. the differences in each fitness measure are below a certain threshold) the current
15 current solution could be selected if it is better than the current solution on three out of the four measures (i.e. on a majority of the measures).

 However, even if the logical comparison is made more sophisticated in this way, it is difficult to realise a flexible comparison, capable of giving
20 different weights to the different performance aspects, using this kind of logical comparison.

 In the second embodiment of the present invention, a unified measure of fitness is calculated for each candidate solution, and the respective unified measures
25 for the different candidate solutions are compared to select the best solution.

 Before describing the process performed in the second embodiment, an explanation of a preferred unified fitness measure will be described with
30 reference to Figs. 10 to 14. This preferred unified measure of fitness in the second embodiment is based on a product of the individual fitness measures for the candidate solution concerned. To facilitate the explanation, in these figures only three of the four
35 fitness measures are represented, namely the first, third and fourth measures described above. This

enables each set of fitness measures to be represented as a plane in three-dimensional space having dimensions of Eff_Oct, Delay_Profile and Fairness.

In Fig. 10, a first candidate solution η_1 has a first set of values E_1 , D_1 and F_1 for the three fitness measures. A second candidate solution η_2 has a second set of fitness values E_2 , D_2 and F_2 . In the case illustrated in Fig. 10, $E_1 > E_2$, $D_1 > D_2$ and $F_1 > F_2$ and it is absolutely clear that the solution η_1 having the higher plane is the better solution. The logical comparison scheme of the first embodiment copes with this kind of scenario well.

A more complicated scenario is illustrated in Fig. 11. In this case, if the unified fitness measure is based on a product of the individual fitness measures, there is a risk that one solution which has a much higher value of a particular fitness measure than all the other solutions will dominate, even if in terms of the remaining fitness values it is a worse solution. As shown in Fig. 11, the solution η_1 has a much higher fairness value F_1 than that of the solution η_2 . The remaining fitness values for solution η_1 are worse than for the solution η_2 . However, the product $F_1.E_1.D_1$ of the fitness measures for solution η_1 is greater than the product $F_2.E_2.D_2$ of the fitness measures for solution η_2 , so the solution η_1 dominates. Although a situation such as that shown in Fig. 11 can always arise, there is less chance of it arising as the number of candidate solutions increases. Accordingly, in embodiments of the present invention it is desirable to have as many iterations per scheduling instant as can be performed within the available processing time.

Fig. 12 shows another problem (mismatch) which needs to be addressed when designing the unified fitness measure. In this figure, even for any particular solution one of the fitness measures has a

(
much higher value than each of the other fitness measures. In this case, any product-based unified measure will place undue emphasis on the single fitness measure having the highest value, with the risk that performance in terms of the other fitness measures is neglected.

In order to try to address this mismatch problem, it is preferable to apply a mapping function to one or more of the individual fitness measures so that after mapping the fitness measures are of the same order as illustrated in Fig. 13. The mapping function is, for example, a normalising function. For example, each fitness measure may be scaled relative to a peak value of the measure identified from previous outcomes.

It is also preferable that the unified fitness measure be controllable or tunable to create different scenarios. This can be achieved by weighting the different individual fitness measures of a candidate solution using controllable weighting coefficients. An extreme example of such a weighting process is illustrated in Fig. 14. In this example, the throughput measure `Eff_Oct` has a weighting coefficient of zero, with the result that the throughput fitness measures of all candidate solutions have the same value, for example one. The remaining fitness measures have non-zero weighting coefficients so that these other measures exclusively influence the unified fitness measure. Accordingly, the solutions with maximum fairness and minimum delay will tend to be selected in this scenario.

Taking the matters set out above into account, a preferred function for determining the unified measure of fitness, for the candidate solution η , i.e. fitness function $f(\eta)$, is

$$f(\eta) = \frac{W_{1E} + W_{2E} \cdot u(\text{Eff_Oct})}{W_{1E}} \cdot \frac{W_{1D} + W_{2D} \cdot v(\text{Delay_Profile})}{W_{1D}} \cdot \frac{W_{1F} + W_{2F} \cdot x(\text{Fairness})}{W_{1F}} \cdot \frac{W_{1R} + W_{2R} \cdot y(\text{Ratio_Waiting_Oct})}{W_{1R}} \quad (25)$$

where $u(\cdot)$, $v(\cdot)$, $x(\cdot)$ and $y(\cdot)$ are the mapping functions, W_{1E} , W_{2E} , W_{1F} , W_{2F} , W_{1D} , W_{2D} , W_{1R} and W_{2R} are the weighting coefficients.

5 The second embodiment tries to maximise the fitness metric considering all the QoS parameters in (1). By employing a tuning mechanism, it reaches a high level of flexibility and control over different parameters such as fairness, total throughput, average
10 throughput and total delay.

Conventional scheduling techniques such as RR or Max C/I in a way try to perform this maximisation process per TTI. However RR scheduling, for example, puts most effort on Fu in (1) by ignoring the other
15 parameters such as Th to some extent while max C/I scheduling mostly concentrates on Th and De .

The processing in the second embodiment will now be described with reference to Fig. 15.

Initially a decision is made based on the selected
20 policy, about the values to be assigned to the weighting coefficients. For example in (25), if just the fairness matters W_{1F} and W_{2F} is set to one, W_{2E} , W_{2D} , W_{2R} are set to small values and W_{1E} , W_{1D} , W_{1R} are set to very large values, the algorithm tends to behave
25 similar to RR.

The processing in Fig. 15 is the same as in the first embodiment (Fig. 9) except for steps S3 and S5. In the second embodiment, in a step S3' a unified fitness measure $f(\eta_{C/I})$ for the max C/I solution is
30 calculated using (25). In step S5, a unified fitness measure $f(\eta_{\text{current}})$ is calculated using (25).

Fig. 16 is a flowchart for use in explaining the processing carried out by the second generator 124 in Fig. 4.

In a first step S20 processing starts for the scheduling instant (TTI) under consideration.

The generator 124 maintains a list of UEs to which data is waiting at the transmitter to be transmitted. A UE has waiting data if its source queue is non-empty. An example of the waiting list is shown in Fig. 17(a). In step S21 the UE at the bottom of the list is moved to the top. In the present example, UE34 down at the bottom of the list in Fig. 17(a) is moved to the top of the list. All other UEs are moved down one place in the list.

In step S22 the list is updated by adding any new UEs which now have waiting data and deleting any existing UEs for which there is no longer any waiting data. Thus, after steps S21 and S22 the list may appear as shown in Fig. 17(b). UE49 has been deleted as it no longer has any waiting data, and a new UE, UE30, is added to the bottom of the list.

In the simplest RR solution the UE with the highest priority (UE34 in Fig. 17(b)) is allocated all the available channels (channels 1 to 10). In step S23 a measure of throughput is calculated for each available MCS level (and at the particular power level chosen). Also, a C/I value is obtained from a C/I report received from the specified UE (i.e. UE34 in this example). Then in step S24 the generator 124 decides the most efficient MCS level for each channel. After this, the amount of data (number of octets) to be transmitted on each channel is decided.

In step S25 all the available channels are assigned to the selected UE (the UE at the top of the list).

In step S26 the processing for the scheduling instant under consideration is terminated and the solution η_{RR} is output.

Fig. 18 shows a variation of the basic RR scheduling process which is referred to as a packet-weighted RR process. Steps S20 to S22 are the same as in Fig. 16.

In step S30, for each UE in the list, a measure of throughput is calculated for each available MCS level. Also, a C/I report is obtained from each UE in the list.

In step S31 the UE at the top of the list is selected as the current UE to be processed. In this case, however, the UE at the top of the list is not the only UE which will be allocated channels. Thus, an assessment is made in step S31 of the source queue fill level for the current UE in relation to the other UEs in the list. Then, in step S31, the number of channels to be allocated to the current UE, the MCS level and the number of octets to be transmitted on each channel allocated to the current UE are decided.

In step S32 the determined number of channels is assigned to the selected user. For example in Fig. 19(a) the UE11 is assigned channels 1 and 2. The minimum number of channels assignable is 1. In step S33 the number of channels remaining for allocation is determined. In this example, after allocating channels 1 and 2 to UE11, eight channels (channels 3 to 10) remain.

In step S34 it is checked whether there is any remaining available channel. If there is, then in step S36 the next UE in the list is selected (UE5). The process of steps S31 to S35 is repeated for this UE. Thus, as shown in Fig. 19(a) UE5 is allocated channel 3. The processing continues in this way until all the available channels have been allocated. Eventually,

after allocating the last channel (channel 10) to UE8, there is no channel remaining. The remaining UEs do not get allocated channels in this scheduling instant. In this example, UEs 23 and 34 are not allocated

5 channels.

Fig. 20 shows an example of the process performed by the first generator 122 in Fig. 4 to generate the max C/I candidate solution $\eta_{C/I}$.

10 In step S20 processing begins for the scheduling instant under consideration. In step S40 the list of UEs with waiting data is updated in the same way as in the previously-described step S22 of the RR process. After this, in step S42, the list is reordered in order of decreasing reported C/I value.

15 The following steps S44 to S47 correspond respectively to the previously-described steps S23 to S26 in the RR process of Fig. 16. Thus, in this process all the channels are allocated to just the single UE at the top of the list, i.e. the UE with the
20 best reported C/I value.

It is also possible for the first generator to use a packet-weighted variant of the process, as shown in Fig. 21.

25 This process is similar to the basic max C/I process of Fig. 20. However, instead of selecting a single UE at the top of the list and allocating all the available channels to that one UE, the packet-weighted process allocates channels to two or more UEs occupying the highest positions in the list. This is done by
30 first applying the step S50 and then applying steps S51 to S56 iteratively. Steps S50 to S56 correspond respectively to steps S30 to S36 in the packet-weighted RR process of Fig. 18.

Incidentally, as is known in the wireless
35 communication art, it is possible for a UE to have two or more possible channels available for use in the

(
downlink direction. These channels may have different carrier frequencies, for example. In this case, the UE produces C/I reports for each available channel, and the base station and/or the mobile station selects one
5 of the available channels (carriers). This may be taken into account by a scheduler embodying the present invention, so that the selection of the available channel is part of each candidate solution.

Next, simulation results for embodiments of the
10 invention are provided. These simulation results relate to simulated worldwide web (WWW) browsing sessions. The simulations are based on parameters described in TR25.848, ver1.0.0, RP-010191, TSG-RAN #11 March 2001. A WWW browsing session comprises a
15 sequence of packet calls. A UE initiates a packet call when requesting an information entity. In the present simulation, a TTI is assumed to be of duration 2ms (as currently specified in the 3GPP standard). The concept of "efficient throughput" for a service is defined as
20 the portion of the throughput which satisfies the QoS requirements within the specified delay threshold γ . In the present simulation it is assumed that the acceptable delay threshold γ (delivery time delay) is 1.5 seconds. Chase combining is carried out and a
25 maximum of six retransmissions is permitted. A packet is dropped if it cannot be delivered within six retransmissions.

It is also assumed that the channel estimation performed by each UE is perfect, and that the feedback
30 signalling is error-free. A minimum reporting delay is considered to be three TTIs. This represents the minimum time between the transmission of a packet and the receipt of an acknowledgement message ACK from the UE. The simulation was carried out for a period of ten
35 seconds (5000 TTIs and for 50 UEs.

Fig. 22 shows the simulated packet data arriving in the source queues at the transmitter base station for the 50 UEs.

For each UE a simulated C/I scenario is generated, for example using a Gaussian autoregressive model. In order to simulate extreme conditions the UEs are assumed to have significant C/I differences. As shown in Fig. 23, the C/I value is the same for any given UE over all TTIs. The fixed C/I value increases progressively from -2dB for UE50 to +5dB for UE1. Accordingly, the best C/I value belongs to UE1 and the worst C/I value belongs to UE50.

Fig. 24 shows a histogram representing packet data activity in the second embodiment of the present invention described above with reference to Fig. 15. Normalisation as described with reference to Fig. 13 is applied.

For each UE three parameters are plotted. The first parameter, indicated by a dotted line, represents the total number of packets arriving at the base station (node B) for the UE concerned. The second parameter, indicated by a solid line, represents the total number of packets transmitted (including retransmissions) to the UE concerned. The third parameter, indicated by a dashed line, represents the total number of packets successfully delivered to the UE concerned.

Fig. 25 shows a corresponding histogram of packet data activity for a max C/I scheduler. Comparing the histograms of Figs. 24 and 25, it is apparent that although the max C/I scheduler achieves a higher overall throughput than a scheduler embodying the present invention, it completely fails to deliver even a single packet to a certain UEs 41 to 44 and 47 to 50. Thus, a scheduler embodying the present invention can

achieve a greater degree of fairness than the max C/I scheduler.

Fig. 26 is a graph for comparing average throughput versus delay in a scheduler embodying the present invention (solid line) and in the max C/I scheduler (dotted line). The lines indicate a mean cumulative density function (CDF) of the throughputs of all UEs.

Fig. 27 shows another graph for comparing overall throughput versus delay in a scheduler embodying the present invention and in the max C/I scheduler.

The results in Fig. 26 demonstrate that a scheduler embodying the present invention manages to outperform the max C/I scheduler in terms of the average delivered throughput. In other words the QoS (efficient throughput) of a scheduler embodying the present invention is 0.03 better than that of the max C/I scheduler, i.e. a scheduler embodying the invention delivers 0.03 more packets within the acceptable threshold of 1.5 seconds for the simulated heavily-loaded scenario with WWW sessions involving 50 users.

Fig. 27 demonstrates that, although the overall throughput of the max C/I scheduler is better for higher delay values (1.1 seconds and above), for smaller and (in most applications critical) delay values, a scheduler embodying the present invention outperforms the max C/I scheduler in terms of overall delivered throughput.

By adjusting the tuning parameters, a scheduler embodying the present invention is also able to achieve any desired combination of throughput, delay, fairness and priorities of various services.

Fig. 28 is a histogram representing a distribution of the delivered packets in time in a scheduler embodying the present invention, and Fig. 29 is a comparable histogram for the max C/I scheduler. As can

(
be seen, for the max C/I scheduler there is a concentration of packets delivered with delays of around 4.0 seconds. A scheduler embodying the present invention, on the other hand, manages to deliver the
5 majority of the packets with much lower delays, for example within a delay of 1.0 seconds.

Accordingly, a scheduler embodying the present invention can outperform the conventional scheduling techniques such as max C/I and RR considerably. A
10 scheduler embodying the present invention can also achieve a high level of tuning ability and flexibility to control various system parameters.

Although an example of the present invention has been described above in relation to a wideband CDMA
15 network having an asynchronous packet mode, it will be appreciated that the present invention can also be applied to any other networks in which a scheduling problem arises. These networks could be, or could be adapted from, other CDMA networks such as an IS95
20 network. These networks could also be, or be adapted from other mobile communication networks not using CDMA, for example networks using one or more of the following multiple-access techniques: time-division multiple access (TDMA), wavelength-division multiple
25 access (WDMA), frequency-division multiple access (FDMA) and space-division multiple access (SDMA).

Although embodiments of the present invention have been described as having distinct "units", those skilled in the art will appreciate that a
30 microprocessor or digital signal processor (DSP) may be used in practice to implement some or all of the functions of the base station and/or mobile station in embodiments of the present invention.

CLAIMS:

1. A packet scheduling method, for scheduling packets of data for transmission from a transmitter via at least one channel to a plurality of receivers, which
5 method comprises:

generating a plurality of candidate scheduling solutions, at least one said candidate solution being generated randomly, and each candidate solution specifying at least the receiver(s) to which packets
10 are to be transmitted in a scheduling instant under consideration; and

comparing the generated candidate solutions and, based on the comparison results, selecting a best one of the candidate solutions to use to transmit packets
15 in said scheduling instant.

2. A method as claimed in claim 1, wherein at least one of the candidate solutions is generated using a deterministic scheduling method.

3. A method as claimed in claim 2, wherein one
20 said candidate solution is a round-robin scheduling solution generated by a round-robin scheduling method.

4. A method as claimed in claim 2 or 3, wherein one said candidate solution is a maximum carrier-to-interference ratio solution generated by a maximum
25 carrier-to-interference ratio scheduling method.

5. A method as claimed in any preceding claim, wherein a plurality of channels are available for transmitting packets from the transmitter to the receivers in each scheduling instant, and each said
30 candidate solution further specifies one or more parameters applicable to each said available channel in the scheduling instant under consideration.

6. A method as claimed in claim 5, wherein each said candidate solution specifies how the specified
35 receiver(s) is (are) to be allocated to the available channels in the scheduling instant under consideration.

(
7. A method as claimed in claim 5 or 6, wherein each said candidate solution specifies a modulation and/or coding scheme to be applied to each said available channel in the scheduling instant under
5 consideration.

8. A method as claimed in any one of claims 5 to 7, wherein each said candidate solution specifies a transmission power for each said available channel in the scheduling instant under consideration.

10 9. A method as claimed in any one of claims 5 to 8, wherein each said candidate solution further specifies an amount of data to be transmitted via each said available channel in the scheduling instant under consideration.

15 10. A method as claimed in any preceding claim, carried out over a series of iterations, wherein in each said iteration a new candidate solution is generated and compared with a candidate solution selected in a previous iteration as said best solution,
20 the new candidate solution being selected as a new best solution if, based on the comparison result, it is judged to be a better solution than the previously-selected best solution.

25 11. A method as claimed in claim 10, wherein the plurality of candidate solutions include a round-robin solution determined using a round-robin scheduling method, and a maximum carrier-to-interference ratio solution determined using a maximum carrier-to-interference ratio scheduling method, said maximum
30 carrier-to-interference solution being generated before an iteration in which the round-robin solution is generated.

35 12. A method as claimed in claim 11, wherein said maximum carrier-to-interference ratio solution is generated in an initialisation step carried out before the first iteration of said series of iterations, that

(
maximum carrier-to-interference ratio solution being set as the initial best solution against which the new candidate solution generated in the first iteration is compared.

5 13. A method as claimed in claim 12, wherein said round-robin solution is the new candidate solution generated in the first iteration.

14. A method as claimed in claim 13, wherein the new candidate solutions generated in the second and
10 each subsequent iteration of the series are randomly-generated solutions.

15. A method as claimed in any preceding claim, wherein, for each said candidate solution, at least one fitness measure is determined, the or each said fitness
15 measure being a measure of performance of the candidate solution concerned, and said best solution is identified based on a comparison of the fitness measure(s) of the generated candidate solutions.

16. A method as claimed in claim 15, wherein a
20 set of two or more such individual fitness measures is determined for each said candidate solution, which fitness measures relate to different respective performance aspects.

17. A method as claimed in claim 16, wherein, for
25 each said performance aspect, the corresponding individual fitness measures for the different candidate solutions are compared to select the best solution.

18. A method as claimed in claim 17, wherein the best solution is the solution which has the best
30 fitness measure in every one of said performance aspects.

19. A method as claimed in claim 16, wherein said two or more individual fitness measures in said set for each candidate solution are combined together using a
35 fitness function to produce a unified measure of fitness for the candidate solution concerned, and the

(
respective unified measures of fitness for the candidate solutions are compared to select said best solution.

20. A method as claimed in claim 19, wherein at least one said individual fitness measure of said set is adjusted to reduce a mismatch between it and another one of the individual fitness measures of said set.

21. A method as claimed in claim 19 or 20, wherein, for one or more said individual fitness measures of said set, a mapping function is applied to the fitness measure so as to reduce a mismatch between it and another one of said individual fitness measures of said set.

22. A method as claimed in any one of claims 19 to 21, wherein one or more of the individual fitness measures of said set are normalised.

23. A method as claimed in any one of claims 19 to 22, wherein one or more of the individual fitness measures of said set are weighted when the individual fitness measures are combined together to produce said unified measure of fitness.

24. A method as claimed in any one of claims 19 to 23, wherein said unified measure of fitness is produced by forming a product of said individual fitness measures in said set.

25. A method as claimed in any one of claims 15 to 24, wherein one said fitness measure is a measure of total data throughput to the specified receiver(s) for the candidate solution concerned.

26. A method as claimed in any one of claims 15 to 25, wherein one said fitness measure is a measure of an amount of data at the transmitter waiting for transmission to the specified receiver(s) for the candidate solution concerned.

27. A method as claimed in any one of claims 15 to 26, wherein one said fitness measure is a measure of

(
delay(s) in transmitting data to the specified receiver(s) in the candidate solution concerned.

28. A method as claimed in any one of claims 15 to 27, wherein one said fitness measure is a measure of how said candidate solution affects a fairness of said packet scheduling amongst active ones of said plurality of receivers.

29. A method as claimed in claim 27, wherein the or each said delay is calculated relative to a preselected delay parameter applicable to a type of data to be transmitted to the receiver concerned.

30. A method as claimed in claim 29, wherein the delay parameter applicable to a higher-priority type of data is lower than the delay parameter applicable to a lower-priority type of data.

31. A method as claimed in claim 25, wherein said measure of total throughput is calculated by estimating a total amount of data which will be transmitted successfully to the specified receiver(s) of the candidate solution in the scheduling instant under consideration.

32. A method as claimed in claim 26, wherein said measure of waiting data is a ratio of a current amount of data waiting for transmission at said transmitter to each specified receiver to a total the amount of data which has arrived at said transmitter for transmission to the receiver concerned.

33. A method as claimed in claim 27, wherein said measure of delay is calculated based on a sum of respective delay times for transmissions to the specified receiver(s), each said delay time being dependent upon a difference between a current time and a time of arrival at said transmitter of the earliest data currently waiting for transmission to the specified receiver.

34. A method as claimed in claim 28, wherein said fairness measure is based on a variance or average of respective estimated data throughputs to all active ones of said receivers if said candidate solution is used.

35. A method as claimed in any preceding claim, wherein said transmission is a wireless transmission.

36. A method as claimed in any preceding claim, wherein said transmitter is part of a base station of a wireless communication system, and each said receiver is part of a user equipment of that system.

37. Packet scheduling apparatus, for scheduling packets of data for transmission from a transmitter to a plurality of receivers via at least one channel, which apparatus comprises:

solution generating means for generating a plurality of candidate scheduling solutions, at least one said candidate solution being generated randomly, and each said candidate solution specifying at least the receiver(s) to which packets are to be transmitted in a scheduling instant under consideration; and

best solution selecting means for comparing the generated candidate solutions and selecting, based on the comparison results, a best one of the candidate solutions to use for transmitting packets in the scheduling instant.

38. A transmitter comprising:

packet scheduling apparatus as claimed in claim 37; and

transmitting means connected operatively to said packet scheduling apparatus and operable, when scheduling instant occurs, to cause packets to be transmitted to the receiver(s) specified by the best solution selected for that instant by the packet scheduling apparatus.

(39. A packet scheduling method substantially as hereinbefore described with reference to any of Figs. 4 to 29 of the accompanying drawings.

40. Packet scheduling apparatus substantially as
5 hereinbefore described with reference to any of Figs. 4 to 29 of the accompanying drawings.

41. A transmitter substantially as hereinbefore described with reference to any of Figs. 4 to 29 of the accompanying drawings.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



INVESTOR IN PEOPLE

Application No: GB 0216239.4
Claims searched: 1-38

Examiner: Anita Keogh
Date of search: 13 January 2003

Patents Act 1977 : Search Report under Section 17

Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance
A	-	GB 2361147 A (MOTOROLA) see abstract and page 3 line 30 to page 4 line 20
A, E	-	EP 1227626 A2 (MOTOROLA) see abstract and page 2 lines 49-55
A	-	WO 2001/63855 A1 (ERICSSON) see abstract
A	-	CN 12347253 A (YANG) see online abstract

Categories:

X Document indicating lack of novelty or inventive step	A Document indicating technological background and/or state of the art.
Y Document indicating lack of inventive step if combined with one or more other documents of same category.	P Document published on or after the declared priority date but before the filing date of this invention.
& Member of the same patent family	E Patent document published on or after, but with priority date earlier than, the filing date of this application.

Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC^v:

H4K, H4L

Worldwide search of patent documents classified in the following areas of the IPC⁷:

H04L, H04Q

The following online and other databases have been used in the preparation of this search report:

EPODOC, WPI, JAPIO

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)